

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1336.46 Seconds

(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819B-40

Perfect score: 21

Sequence: 1 cccgttggtctacgcgtctg 21

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

FST:\*

1: em\_estba:\*

2: em\_estum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gssl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	147	10	BE140679
2	21	100.0	159	14	T91602
3	21	100.0	229	10	BE140710
4	21	100.0	230	13	C03623

C 78	21	100.0	720	10	BF976100	602244936	C 151	21	100.0	1055	14	CD517144	AGENCOURT
C 79	21	100.0	724	10	BE439689	HTML-360F	C 152	21	100.0	1061	12	BQ054682	AGENCOURT
C 80	21	100.0	732	12	BG699921	602681285	C 153	21	100.0	1083	13	BU594971	AGENCOURT
C 81	21	100.0	732	14	CB997879	AGENCOURT	C 154	21	100.0	1108	13	BQ072806	AGENCOURT
C 82	21	100.0	734	9	AV733676	AV733676	C 155	21	100.0	1111	14	CD517173	AGENCOURT
C 83	21	100.0	735	12	BI762063	603049235	C 156	21	100.0	1201	13	BX363746	AGENCOURT
C 84	21	100.0	743	14	CB956955	AGENCOURT	C 157	21	100.0	1217	12	BMS44066	AGENCOURT
C 85	21	100.0	743	14	CD104670	AGENCOURT	C 158	21	100.0	1722	13	BQ057412	AGENCOURT
C 86	21	100.0	744	12	BI597401	603243611	C 159	20	95.2	338	9	AA323639	AGENCOURT
C 87	21	100.0	745	14	CB984378	AGENCOURT	C 160	19.4	92.4	256	10	BE159476	AGENCOURT
C 88	21	100.0	748	10	BF129007	601811329	C 161	19.4	92.4	405	9	AA244273	AGENCOURT
C 89	21	100.0	751	12	BG759996	602733419	C 162	19.4	92.4	448	10	BF834707	AGENCOURT
C 90	21	100.0	758	14	CB985916	AGENCOURT	C 163	19.4	92.4	464	14	CD705070	AGENCOURT
C 91	21	100.0	761	12	BG541682	602571259	C 164	19.4	92.4	464	14	CD698716	AGENCOURT
C 92	21	100.0	766	12	BG570800	602591286	C 165	19.4	92.4	511	14	CD707609	AGENCOURT
C 93	21	100.0	776	12	BI227278	602948666	C 166	19.4	92.4	533	14	CD687511	AGENCOURT
C 94	21	100.0	789	12	BG538956	602568315	C 167	19.4	92.4	534	12	BG756165	AGENCOURT
C 95	21	100.0	790	12	BI837330	603090674	C 168	19.4	92.4	546	10	AW351777	AGENCOURT
C 96	21	100.0	791	13	BU596174	AGENCOURT	C 169	19.4	92.4	591	14	CB551184	AGENCOURT
C 97	21	100.0	796	12	BG399870	602441384	C 170	19.4	92.4	602	14	CD699993	AGENCOURT
C 98	21	100.0	796	14	CB997069	AGENCOURT	C 171	19.4	92.4	623	14	CD767511	AGENCOURT
C 99	21	100.0	802	10	BE238242	601811992	C 172	19.4	92.4	632	14	CB550782	AGENCOURT
C 100	21	100.0	802	12	BG757550	602714723	C 173	19.4	92.4	645	14	CB551104	AGENCOURT
C 101	21	100.0	806	14	CB993192	AGENCOURT	C 174	19.4	92.4	663	14	CD701461	AGENCOURT
C 102	21	100.0	809	14	CB992720	AGENCOURT	C 175	19.4	92.4	671	12	BG540027	AGENCOURT
C 103	21	100.0	811	12	BI766898	603053138	C 176	19.4	92.4	681	14	CB554193	AGENCOURT
C 104	21	100.0	821	12	BI668794	603234742	C 177	19.4	92.4	693	12	BG685335	AGENCOURT
C 105	21	100.0	823	14	CB993431	AGENCOURT	C 178	19.4	92.4	704	9	AV707666	AGENCOURT
C 106	21	100.0	826	12	BG431923	602498354	C 179	19.4	92.4	708	14	N25099	AGENCOURT
C 107	21	100.0	830	14	CB994002	AGENCOURT	C 180	19.4	92.4	735	12	BG484403	AGENCOURT
C 108	21	100.0	831	10	BF525771	602069914							
C 109	21	100.0	832	14	CB958376	AGENCOURT							
C 110	21	100.0	837	12	BI517599	603042205							
C 111	21	100.0	843	12	BI261592	602953662							
C 112	21	100.0	851	14	CD558780	AGENCOURT							
C 113	21	100.0	854	12	BG756171	602713472							
C 114	21	100.0	857	13	BQ890384	AGENCOURT							
C 115	21	100.0	859	12	BG546165	603573582							
C 116	21	100.0	860	12	BG755507	602713882							
C 117	21	100.0	871	12	BI767351	AGENCOURT							
C 118	21	100.0	872	12	BQ423978	AGENCOURT							
C 119	21	100.0	872	13	BQ721555	AGENCOURT							
C 120	21	100.0	873	12	BG540219	AGENCOURT							
C 121	21	100.0	878	14	CB986657	AGENCOURT							
C 122	21	100.0	880	9	AU139061	AU139061							
C 123	21	100.0	881	14	CD558622	AGENCOURT							
C 124	21	100.0	884	12	BI545349	603187481							
C 125	21	100.0	889	14	CD516536	AGENCOURT							
C 126	21	100.0	890	12	BG754449	602710081							
C 127	21	100.0	901	13	BQ955105	AGENCOURT							
C 128	21	100.0	911	10	B8878872	601493167							
C 129	21	100.0	915	10	BF974675	602445314							
C 130	21	100.0	917	12	BG397584	602438625							
C 131	21	100.0	923	13	BQ642494	AGENCOURT							
C 132	21	100.0	926	13	BQ642196	AGENCOURT							
C 133	21	100.0	943	13	BQ060004	AGENCOURT							
C 134	21	100.0	945	12	BG361646	602564242							
C 135	21	100.0	953	10	BF975606	602246064							
C 136	21	100.0	959	10	BF876257	602245136							
C 137	21	100.0	959	13	BQ057749	AGENCOURT							
C 138	21	100.0	960	13	BQ894372	AGENCOURT							
C 139	21	100.0	967	12	BI553587	603190506							
C 140	21	100.0	980	10	BE974318	602243975							
C 141	21	100.0	980	13	BQ056978	AGENCOURT							
C 142	21	100.0	990	12	BQ056978	AGENCOURT							
C 143	21	100.0	992	12	BG686860	602650843							
C 144	21	100.0	1005	13	BG759351	602711866							
C 145	21	100.0	1007	13	BQ062257	AGENCOURT							
C 146	21	100.0	1007	13	BQ057757	AGENCOURT							
C 147	21	100.0	1015	13	BQ060930	AGENCOURT							
C 148	21	100.0	1025	12	BQ054502	AGENCOURT							
C 149	21	100.0	1026	13	BQ061344	AGENCOURT							
C 150	21	100.0	1027	13	BQ064032	AGENCOURT							
	21	100.0	1030	13	BQ063085	AGENCOURT							

## ALIGNMENTS

BE140679 147 bp mRNA linear EST 21-JUN-2000  
 CMO-HT0016-170699-004 HT0016 Homo sapiens cDNA, mRNA sequence.  
 BE140679  
 BE140679.1 GI:8603400  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (Bases 1 to 147)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=CMO-HT0016-170699-004&t3=1999-06-17&t4=1>)  
 Seg primer: puc 18 forward  
 High quality sequence start: 2  
 High quality sequence stop: 147.



## FEATURES

Location/Qualifiers

1. .147  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0016"  
 /notes="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCATGCGTCTG 21

Db 114 CCTGTGCTCATGCGTCTG 134

## RESULT 2

T91602/C

## LOCUS

Y921b03.r1 Stratagene lung (#937210) linear EST 22-MAR-1995  
 IMAGE:118349 5' similar to gb:K01506 HLA CLASS II  
 HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN); mRNA  
 sequence.

## ACCESSION

T91602

## VERSION

EST

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 159)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,  
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
 Roalson, T., Schellenberg, K., Soares, M.B., Tan, P., Thierry-Mieg, J.,  
 Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
 and Marra, M.

## AUTHORS

Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)

## TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 114

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium [info@image.llnl.gov] for further information.

Seq primer: M13RP1

High quality sequence stop: 114.

Location/Qualifiers

1. .159

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0016"

/sex="male"

/dev\_stage="72 years"

/lab\_host="SOLR cells (kanamycin resistant)"

/clone\_lib="Stratagene lung (#937210)"

/note="Organ: lung; Vector: pBluescript SK-; Site 1:  
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
 Oligo dr. normal lung. Average insert size: 1.0 kb;  
 Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGGAG  
 3' ~3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCATGCGTCTG 21

Db 104 CCTGTGCTCATGCGTCTG 84

## RESULT 3

BE140710

## LOCUS

BE140710 229 bp mRNA linear EST 21-JUN-2000  
 CM0-HT0016-140699-008 HT0016 Homo sapiens cDNA, mRNA sequence.

## ACCESSION

BE140710

## VERSION

EST

## KEYWORDS

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 229)

## AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/ILICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=8t2=CM0-HT0016-140

699-008&amp;t3=1999-06-14&amp;t4=1)

Seq primer: puc 18 forward

High quality sequence start: 7

High quality sequence stop: 229.

Location/Qualifiers

1. .229

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0016"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

Query Match 100.0%; Score 21; DB 10; Length 229;

Best Local Similarity 100.0%; Pred. No. 17;

## ORIGIN

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 102 CCTGTGGTCTATGCGTCTG 122

RESULT 4  
 C03623/c  
 LOCUS C03623 Human heart cDNA (YNakamura) linear EST 30-JUN-1996  
 DEFINITION 3NHCI920, mRNA sequence.  
 ACCESSION C03623  
 VERSION C03623.1 GI:1466874  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 230)  
 AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.  
 TITLE Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing  
 JOURNAL Genomics 35 (1), 231-235 (1996)  
 MEDLINE 96299762  
 PUBMED 8661126  
 COMMENT Contact: Yusuke Nakamura  
 Institute of Medical Science  
 University of Tokyo  
 4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan  
 Tel: 81-3-5449-5372  
 Fax: 81-3-5449-5433  
 Email: yusuke@ims.u-tokyo.ac.jp.

FEATURES  
 source  
 1..230  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="3NHCI920"  
 /dev\_stage="adult"  
 /clone\_lib="Human heart cDNA (YNakamura)"  
 /notes="Organ: heart; normalized directionally cloned cDNA from adult heart"

ORIGIN  
 Query Match 100.0%; Score 21; DB 13; Length 230;  
 Best Local Similarity 100.0%; Pred. NC. 17;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 216 CCTGTGGTCTATGCGTCTG 196

RESULT 5  
 CD707370/c  
 LOCUS CD707370 249 bp mRNA linear EST 25-JUN-2003  
 DEFINITION EST23897 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
 ACCESSION CD707370  
 VERSION CD707370.1 GI:32238000  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 249)  
 AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-X., Pan,Z.-G. and Zeng,Y.-X.  
 TITLE Transcriptional Gene Expression Profile of Human Nasopharynx Unpublished (2003)  
 JOURNAL  
 COMMENT Contact: Yixin Zeng  
 Cancer Center

Sun Yat-sen University  
 651 Dongfeng Road East, Guangzhou 510060, China  
 Tel: 86-1380-3770-7433  
 Fax: 86-20-8775-4506  
 Email: yxzeng@gzsums.edu.cn.

FEATURES  
 Location/Qualifiers  
 1..249  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN  
 Query Match 100.0%; Score 21; DB 14; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 237 CCTGTGGTCTATGCGTCTG 217

RESULT 6  
 AW799168/c  
 LOCUS AW799168 263 bp mRNA linear EST 16-MAY-2000  
 DEFINITION RC0-UM0051-c10300-011-c05 UM0051 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW799168  
 VERSION AW799168.1 GI:7851038  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 263)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=tl2=RC0-UM0051-010300-011-c05&tl3=2000-03-01&tl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 29  
 High quality sequence stop: 263.  
 Location/Qualifiers  
 1..263  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="UM0051"  
 /note="Organ: uterus; Vector: puc18; Site 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No.

196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 263;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTGGCTATGCGTCTG 21  
|||||  
Db 49 CCCTGTGGCTATGCGTCTG 29

## RESULT 7

BE242285/c

## LOCUS

## DEFINITION

BE242285 265 bp mRNA linear EST 03-OCT-2001  
TCAAP1562 Pediatric acute myelogenous leukemia cell (FAB M1)  
Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1562, mRNA  
sequence.

## ACCESSION

BE242285

## KEYWORDS

## SOURCE

EST.

BE242285.1 GI:9094012

## ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 265)

Wei Y., Tsang Y.T.M., Mei, G., Ku, J.X., Ali-Osman Jr., F.R.,

Muzny, D., Soucek, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project

Unpublished (2000)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@tccc.org

Citation: Carrincci, P. and Hayashizaki Y. High efficiency

full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M13 primer.

Location/Qualifiers

1..265

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TCAAP1562"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="myeloid cell"

/dev\_stage="pediatric 6 years"

/lab\_host="DH10s"

/clone\_lib="pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project=TCAA"

/note="Vector: lambda pBS; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(AT) primer [5'GGAGGACTCGAGCGCGCAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGAGCTCGATCCGCGCCGCAATATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pBS vector. Library went through one round of

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Carrincci P, Westover A, Nishiyama Y, Ohsumi T,

Itoh M, Nagaoaka S, Sasaki, Okazaki Y, Muramatsu M,

Schneider C, Hayashizaki Y, High efficiency selection of

full-length cDNA by improved biotinylated cap trapper.,

DNA Res 4: 1, 61-6, Feb 28, 1997)"

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 289;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 21; DB 10; Length 265;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTGGCTATGCGTCTG 21  
|||||  
Db 208 CCCTGTGGCTATGCGTCTG 188

## RESULT 8

AA360953/c

## LOCUS

## DEFINITION

AA360953 289 bp mRNA linear EST 21-APR-1997  
EST70157 T-cell lymphoma Homo sapiens cDNA 5' end similar to major  
histocompatibility complex, class II antigen, alpha chain  
(GB:X03100), mRNA sequence.

## ACCESSION

AA360953

## VERSION

AA360953.1 GI:2013273

## KEYWORDS

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 289)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Feldner, R.A.,

Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,

White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,

Clayton, R.A., Cline, T.R., Cotton, M.D., Barle-Hughes, J., Fine, L.D.,

Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,

Glocke, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,

Kleek, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Scott, J.L., Saudak, D.M., Shirley, R.,

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,

Bednarek, P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, R.J.,

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,

He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,

Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,

Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M.,

Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseitine, W.A., Fields, C.,

Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

Nature 377 (6547 Suppl), 3-174 (1995)

7566098

Other\_ESTs: TNC172266

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..289

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="ATCC (inhost):165076"

/db\_xref="taxon:9606"

/cell\_type="T-lymphocyte"

/clone\_lib="T-cell lymphoma"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:

XhoI"

## ORIGIN

QY 1 CCTGTTGGTCTATGCGTCTG 21  
|||||  
Db 183 CCTGTTGGTCTATGCGTCTG 163

RESULT 9  
CD706205/c  
LOCUS CD706205 290 bp mRNA linear EST 25-JUN-2003  
DEFINITION human nasopharynx Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD706205  
VERSION CD706205.1 GI:32236835  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 290)  
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
JOURNAL Unpublished (2003)  
COMMENT Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@gzsums.edu.cn.

FEATURES  
source  
1..290  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 290;  
Best Local Similarity 100.0%; Freq. No. 19;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTTGGTCTATGCGTCTG 21  
|||||  
Db 221 CCTGTTGGTCTATGCGTCTG 201

RESULT 10  
CD694567/c  
LOCUS CD694567 295 bp mRNA linear EST 25-JUN-2003  
DEFINITION human nasopharynx Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD694567  
VERSION CD694567.1 GI:32219338  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 295)  
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
JOURNAL Unpublished (2003)  
COMMENT Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@gzsums.edu.cn.

FEATURES  
source  
1..295  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S1SNUS-8-C07"  
/sex="F"  
/tissue\_type="Asciites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-5"  
/lab\_host="Top10F"  
/clone\_lib="S1SNUS"  
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of 74 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 295;  
Best Local Similarity 100.0%; Freq. No. 19;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTTGGTCTATGCGTCTG 21  
|||||  
Db 212 CCTGTTGGTCTATGCGTCTG 192

RESULT 11  
BM772902/c  
LOCUS BM772902 335 bp mRNA linear EST 04-MAR-2002  
DEFINITION K-EST0057161 S1SNUS Homo sapiens cDNA clone S1SNUS-8-C07 5', mRNA sequence.  
ACCESSION BM772902  
VERSION BM772902.1 GI:19102517  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 335)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,C.M., Park,H.S., Kim,S. and Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 8 row: C column: 07  
High quality sequence stop: 335.  
Location/Qualifiers  
1..335  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S1SNUS-8-C07"  
/sex="F"  
/tissue\_type="Asciites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-5"  
/lab\_host="Top10F"  
/clone\_lib="S1SNUS"  
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of 74 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTATGCGTCTG 21  
 |||||  
 Db 232 CCTGTGCTATGCGTCTG 212

## RESULT 12

T94759/3  
 LOCUS T94759.1 GI:728247  
 DEFINITION Homo sapiens (human)  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 366)  
 Chisoe, S., Dietrich, N., Dubuque, T., Favallo, A., Gish, W.,  
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,  
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
 Rohlfing, T., Schellenberg, K., Soares, M. B., Tan, P., Thierry-Mieg, J.,  
 Trevas, R., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,  
 and Marra, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 97044478  
 8889549  
 CONTACT: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert size: 1152  
 High quality sequence stops: 251 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1152 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 251.  
 Location/Qualifiers  
 1..366  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="Gene:488240"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:119951"  
 /sex="male"  
 /dev\_stage="72 years"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Stratagene lung (#937210)"  
 /note="Organ: lung; Vector: pBluescript SK-; Site 1:  
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
 Oligo dt. normal lung. Average insert size: 1.0 kb;  
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGACG  
 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 14

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM&lt;/

```

Email: yxzeng@zsums.edu.cn.
FEATURES
  Source
    Location/Qualifiers
      1..390
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /tissue_type="normal nasopharynx"
        /clone_lib="human nasopharynx"
        /note="ESTs generated from a normal nasopharynx cDNA
        library from southern Chinese"
ORIGIN
  Query Match      100.0%; Score 21; DB 14; Length 390;
  Best Local Similarity 100.0%; Pred. No. 21;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTTGGTCTATCGCTCG 21
    |||||
Db 241 CCTCTTGGTCTATCGCTCG 221
    |||||
RESULT 15
AA838010/c
LOCUS
DEFINITION
  AA838010.1 GI:2913667
  similar to gb:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)
  ALPHA CHAIN (HUMAN), contains Alu repetitive element,, mRNA
  sequence.
ACCESSION
  AA838010
VERSION
  AA838010.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 415)
  Title:
  Journal:
  Authors:
  Comment:
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-romail.nih.gov
  Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: Stratagene, Inc.
  DNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 301.
  Location/Qualifiers
    1..415
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1418797"
      /sex="mixed"
      /tissue_type="colon tumor"
      /lab_host="SOLR (kanamycin resistant)"
      /clone_lib="NCI CGAP Col2"
      /note="Organ: colon; Vector: Bluescript SK-; Site:1;
      EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
      Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
      GAATTCGCGACGAG 3' 3' adaptor sequence: 5'
      CTCGAGTTTITTTTTTTTTTTT 3' Average insert size: 1.2 kb."
ORIGIN
  Query Match      100.0%; Score 21; DB 9; Length 415;
  Best Local Similarity 100.0%; Pred. No. 21;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTTGGTCTATCGCTCG 21
    |||||
Db 142 CCTCTTGGTCTATCGCTCG 122
    |||||
FEATURES
  Source
    Location/Qualifiers
      1..422
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="3NHCI1710"
        /dev_stage="adult"
        /clone_lib="Human heart cDNA (Ynakamura)"
        /note="Organ: heart; normalized directionally cloned cDNA
        from adult heart"
ORIGIN
  Query Match      100.0%; Score 21; DB 13; Length 422;
  Best Local Similarity 100.0%; Pred. No. 22;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTTGGTCTATCGCTCG 21
    |||||
Db 227 CCTCTTGGTCTATCGCTCG 207
    |||||
RESULT 17
CD102141/c
LOCUS
DEFINITION
  CD102141.1 GI:30755315
  AGNCNCOURT 14007825 NIH MGC 186 Homo sapiens cDNA clone
  IMAGE:30370583 5', mRNA sequence.
ACCESSION
  CD102141
VERSION
  CD102141.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 423)
  Title:
  Journal:
  Authors:
  Comment:
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-romail.nih.gov
  Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
  cDNA Library Preparation: CLONTECH Laboratories, Inc.
  Location/Qualifiers
    423 bp mRNA linear EST 15-MAY-2003
    AGNCNCOURT 14007825 NIH MGC 186 Homo sapiens cDNA clone
    IMAGE:30370583 5', mRNA sequence.
    CD102141
    CD102141.1 GI:30755315
    EST.
    Homo sapiens (human)
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    NIH-MGC http://mgc.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-romail.nih.gov
    Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
    cDNA Library Preparation: CLONTECH Laboratories, Inc.

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: NDCM159 row: m column: 24  
 High quality sequence stop: 417.  
 Location/Qualifiers

## FEATURES

source

1. .423

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="WAGR:30370583"

/lab\_host="DH10B (TI phage-resistant)"

/notes="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI  
 (ggccattatggcc); Site 2: SfiI (ggccgctggcc); Library is  
 oligo-dT primed and directionally cloned. cDNA was  
 prepared from a pooled samples of tissues from Skin,  
 meninges, duramatter, pia matter and choroid plexus. 5'  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCGGCGGCACATG-dT(30)BN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.47 kb (range 0.5C-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH MGC  
 Library"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 423;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTTGGTCTATGGCTCTG 21

|||||

Db 243 CCTGTGTTGGTCTATGGCTCTG 223

## RESULT 18

AW085969

LOCUS

xc76d02.x1 NCI CGAP Ovx32 Homo sapiens cDNA clone IMAGE:2590179 3'  
 similar to gb:U01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)  
 ALPHA CHAIN (HUMAN); mRNA sequence.

ACCESSION

AW085969

VERSION

AW085969.1 GI:6041175

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 427)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb@mail.nih.gov  
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuqui,  
 M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium, LLNL  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 421.

## FEATURES

source

1. .427

/organism="Homo sapiens"

Location/Qualifiers

/mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2590179"  
 /sex="female"  
 /tissue\_type="papillary serous carcinoma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP\_Ov32"

/note="Organ: ovary; Vector: pAMP1; mRNA made from ovarian  
 carcinoma, cDNA made by oligo-dT priming.  
 Non-directionally cloned. Size-selected on agarose gel,  
 average insert size 500 bp. Non-amplified library."

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 427;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTTGGTCTATGGCTCTG 21

|||||

Db 283 CCTGTGTTGGTCTATGGCTCTG 303

## RESULT 19

AW449605

LOCUS

UI-H-B13-aku-g-06-0-UI.s1 NCI CGAP\_Sub5 Homo sapiens cDNA clone  
 IMAGE:2735794 3', mRNA sequence.

ACCESSION

AW449605

VERSION

AW449605.1 GI:6990311

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 427)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb@mail.nih.gov  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:  
 NCI-CGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

1. .427

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2735794"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP\_Sub5"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not 1; Site 2: Eco RI; NCI CGAP\_Sub5  
 is a subtracted library derived from NCI CGAP\_Sub4. The  
 NCI CGAP\_Sub5 library had 3 million recombinants. A  
 single-stranded DNA preparation of NCI CGAP\_Sub4 was used  
 as a tracer in a subtractive hybridization with a driver  
 comprising: the IMAGE pool (NCI CGAP\_Kid3 pool 1 LLAM  
 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonids  
 1322376-1323911, 1456008-1456775, 1500552-1502855);  
 NCI CGAP\_Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778  
 (IMAGE Clonids 1323912-1325831, 1471368-1472903,  
 1492104-1493255); NCI CGAP\_Lus pool 1 LLAM 3575-3582,  
 3851-3854 (IMAGE Clonids  
 -414920-1417991, 1520904-1522439); NCI CGAP\_GC4 pool 1 LLAM  
 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonids

1257096-1258631, 1469064-1470983, 1475592-1476743);  
 NCI\_CGAP\_P222 pool 1 LMAX 2457-2459, 2758-2759, 3062-3068  
 (IMAGE Cloneds 985608-986759, 1101192-1101959,  
 1217928-1220615); NCI\_CGAP\_Co10 pool 1 LMAX 2644-2653,  
 2871-2872 (IMAGE Cloneds  
 1057416-1061255, 1144584-1145351). (10% of the driver  
 population), plus a pool of 3,840 arrayed clones from  
 NCI\_CGAP\_Sub1 (IMAGE Cloneds 2708616-2710535) and  
 NCI\_CGAP\_Sub2 (IMAGE Cloneds 2710536-2712455) (10% of  
 the driver population), plus a pool of 11,136 clones from  
 NCI\_CGAP\_Sub3 (IMAGE Cloneds 2712456-2723591) (10% of the  
 driver population), plus a pool of 5,472 clones from  
 NCI\_CGAP\_Sub4 (IMAGE Cloneds 2723592-2728969) (70% of the  
 driver population). Subtraction was performed as  
 previously described [Bonaldo, Lemon & Soares (1996):  
 Normalization and Subtraction: Two Approaches to  
 Facilitate Gene Discovery. Genome Research 6, 791-806.  
 TAG TISSUE=kidney  
 TAG LIB=NCI CGAP\_Kids  
 TAG\_SEQ=ATTC"

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21

Db 397 CCTGTGGTCTATGCGTCTG 417

## RESULT 20

BI710946/c

LOCUS

DEFINITION id5b07.yl Human insulinoma Homo sapiens cDNA clone IMAGE:5023429  
 5' similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
 ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

ACCESSION

BI710946

VERSION

BI710946.1 GI:15686641

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (Bases 1 to 429)

REFERENCE

AUTHORS

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kastner, K.,  
 Leniska, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Marra, M., Pape, D., Wy-ile, I., Martin, J., Blistain, A.,  
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,  
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R.,  
 Williams, T., Jackson, Y., and Bowers, Y.

TITLE

JOURNAL

COMMENT

Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
 University Genome Sequencing Center For information on obtaining a  
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: 40RP from Gibco

High quality sequence stop: 340.

Location/Qualifiers

1..429

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5023429"

## FEATURES

source

/tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /notes="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
 XhoI; Site\_2: EcoRI; Constructed with lambda ZAPII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/est/lambda\_protocol.shtml).  
 Please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permutt  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21

Db 231 CCTGTGGTCTATGCGTCTG 211

## RESULT 21

BM767805/c

LOCUS

DEFINITION K-EST0050302 SISNU5s2 Homo sapiens cDNA clone SISNU5s2-12-E02 5',  
 mRNA sequence.

ACCESSION

BM767805

VERSION

BM767805.1 GI:19097420

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (Bases 1 to 437)

REFERENCE

AUTHORS

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.R., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 12 row: E column: 02  
 High quality sequence stop: 437.

## FEATURES

source

1..437

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="SISNU5s2-12-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-5"

/lab\_host="Top10"

/clone\_lib="SISNU5s2"

/notes="Organ: Stomach; Vector: pcNS; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dT-selected mRNA by  
 priming with dT-tailed vector. The dT-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was



circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F<sup>+</sup> with electroporation method."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTGGTCTATGCGCTCTG 21

|||||  
233 CCCTGTGGTCTATGCGCTCTG 213

## RESULT 22

BM766631/c

LOCUS

DEFINITION K-EST0048574 S1SNUS52 Homo sapiens cDNA clone S1SNUS52-10-E02 5', mRNA linear EST 04-MAR-2002

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 469)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.F., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE

JOURNAL

COMMENT

21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 10 row: E column: 02

High quality sequence stop: 469.

Location/Qualifiers

1..469

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/c\_name="S1SNUS52-10-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-5"

/lab\_host="Top10F"

/clone\_lib="S1SNUS52"

/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;

Site 2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F<sup>+</sup> with electroporation method."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 469;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTGGTCTATGCGCTCTG 21

|||||  
233 CCCTGTGGTCTATGCGCTCTG 213

## RESULT 23

AV734557/c

LOCUS

DEFINITION AV734557 cda Homo sapiens cDNA clone cGABEB10 5', mRNA linear EST 17-OCT-2000

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 482)

Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,

Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S.,

Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,

Chen,J., Chen,Z. and Han,Z.

Homo sapiens cDNA cda clones

Unpublished (2000)

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..482

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="cGABEB10"

/tissue\_type="pheochromocytoma"

/dev\_stage="Adult"

/lab\_host="BM25.8"

/clone\_lib="cda"

## FEATURES

source

/note="Vector: pT-ripEx2; Site\_1: sfIIA; Site\_2: sfIIB"

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 482;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATCGGCTG 21  
|||||  
DB 200 CCTGTGGTCTATCGGCTG 180

## RESULT 24

LOCUS BF819626 484 bp mRNA linear EST 13-JAN-2001  
DEFINITION MRL-R70028-101100-002-b04 R70028 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF819626  
VERSION BF819626.1 GI:12157598  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (Bases 1 to 484)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.C. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?i=MRL&t2=MRL-R70028-  
101100-002-b04&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 484.

Location/Qualifiers

## FEATURES

source

1..484  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="R70028"

/note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 484;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATCGGCTG 21  
|||||  
DB 406 CCTGTGGTCTATCGGCTG 426

## RESULT 25

BM694247/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..491

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="UI-E-C11-af0-1-16-0-UI"

/tissue\_type="RPE and Choroid"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-C11 is a normalized cDNA library containing the

following tissue(s): RPE and Choroid. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pT7T3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is ACCCTA.

This library was created for the program, Gene Discovery

in the Visual System, supported by National Eye Institute

(NEI)."

(NEI)."

Query Match 100.0%; Score 21; DB 12; Length 491;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATCGGCTG 21

|||||

DB 189 CCTGTGGTCTATCGGCTG 169

## RESULT 26

CB698819/c

```

LOCUS       CD698819              514 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION   EST15342 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD698819
VERSION     CD698819.1   GI:32227504
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 514)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,K., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: fixin zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.

FEATURES             source
            Location/Qualifiers
             1..514
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="normal nasopharynx"
                /clone_lib="human nasopharynx"
                /note="ESTs generated from a normal nasopharynx cDNA
                library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 21; DB 14; Length 514;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  CCCCTGTTGGCTCTATGCGTCTG 21
    |||||
Db   267 CCCCTGTTGGCTCTATGCGTCTG 247

RESULT 27
BM769742/c
LOCUS       BM769742              515 bp      mRNA      linear      EST 04-MAR-2002
DEFINITION   K-BST0053050 S14K402 Homo sapiens cDNA clone S14K402-25-B02 5',
            mRNA sequence.
ACCESSION   BM769742
VERSION     BM769742.1   GI:19099357
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 515)
AUTHORS     Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 25 row: B column: 02
            High quality sequence stop: 515.
            Location/Qualifiers
             1..515
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"

```

```

/clone="S14K402-25-B02"
/cell_line="K402"
/lab_host="Top10F"
/clone_lib="S14K402"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match      100.0%; Score 21; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  CCCGTGTGGCTCTATGCGCTCTG 21
    |||||
Db   210 CCCGTGTGGCTCTATGCGCTCTG 190

RESULT 28
BG541135/c
LOCUS       BG541135              519 bp      mRNA      linear      EST 03-APR-2001
DEFINITION   602569911F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694379 5',
            mRNA sequence.
ACCESSION   BG541135
VERSION     BG541135.1   GI:13533368
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 519)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LIML at:
            http://image.llnl.gov
            Plate: LIM518 row: e column: 04
            High quality sequence stop: 513.
            Location/Qualifiers
             1..519
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4694379"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_lib="NIH_MGC_77"
                /note="Organ: lung; Vector: pDNA-LIB (Clontech); Site 1:
                SfiI (ggccattgcc); Site 2: SfiI (ggccattgcc); 5' and
                3' adaptors were used in cloning as follows: 5' adaptor
                sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
                5'-ATTCTAGAGCCGAGCGGCATG-dT(30)BN-3' (where B = A,
                C, or G and N = A, C, G, or T). Average insert size 1.9
                kb (range 0.5-4.0 kb). 12/15 colonies contained inserts

```

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 519;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATCGCTCTG 21  
|||||  
Db 233 CCTGTGCTCTATCGCTCTG 213

## RESULT 29

CD705087 526 bp mRNA linear EST 25-JUN-2003  
LOCUS EST21614 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
ACCESSION CD705087  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
JOURNAL Unpublished (2003)  
COMMENT Contact: YiXin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@zsums.edu.cn.

FEATURES  
Source  
1..526  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 526;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

QY 1 CCTGTGCTCTATCGCTCTG 21  
|||||  
Db 281 CCTGTGCTCTATCGCTCTG 261

## RESULT 30

CD695435 540 bp mRNA linear EST 25-JUN-2003  
LOCUS EST11958 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
ACCESSION CD695435  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
JOURNAL Unpublished (2003)

## COMMENT

Contact: YiXin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@zsums.edu.cn.

## FEATURES

Source  
1..540  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 540;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATCGCTCTG 21  
|||||  
Db 240 CCTGTGCTCTATCGCTCTG 220

## RESULT 31

BE874055 544 bp mRNA linear EST 20-OCT-2000  
LOCUS 601484433F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3886761 5',  
DEFINITION  
ACCESSION BE874055  
VERSION BE874055.1 GI:10322831  
KEYWORDS  
SOURCE EST.  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DP/Genetics, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM9664 row: b column: 10  
High quality sequence stop: 542.

FEATURES  
Source  
1..544  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:3886761"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_69"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 544;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

VERSION AW406086.1 GI:6925107
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.
FEATURES
     source
     1..547
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:3060714"
         /tissue_type="lymph"
         /cell_type="germinal center B cells"
         /cell_line="MGC85"
         /lab_host="DH10B (LTR)"
         /clone_lib="NIH MGC 37"
         /note="Vector: pVT73-Pac; Site 1: NotI; Site 2: Eco RI;
         Constructed from size fractionated cytoplasmic mRNA
         (1.5-2.5kb). Directionally cloned. Cells provided by Louis
         M. Staudt, Ph.D. Library preparation by Maria de Fatima
         Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 547;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATGCGTCTG 21
|||||
Db 196 CCTGTGGTCTATGCGTCTG 176
|||||

RESULT 35
CD706950/c 547 bp mRNA linear EST 25-JUN-2003
LOCUS human nasopharynx Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION CD706950
VERSION CD706950.1 GI:32237580
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
Zeng, Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
FEATURES
     source
     1..547

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 547;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATGCGTCTG 21
|||||
Db 238 CCTGTGGTCTATGCGTCTG 218
|||||

RESULT 36
BG535978/c 557 bp mRNA linear EST 03-APR-2001
LOCUS NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4688604 5',
DEFINITION mRNA sequence.
ACCESSION BG535978
VERSION BG535978.1 GI:13527523
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1503 row: d column: 13
High quality sequence stop: 557.
FEATURES
     source
     1..557
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:4688604"
         /lab_host="DH10B (T1 phage-resistant)"
         /clone_lib="NIH_MGC_77"
         /note="Organ: lung; Vector: pDNR-LiB (Clontech); Site 1:
         SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and
         3' adaptors were used in cloning as follows: 5' adaptor
         sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence:
         5'-ATTCTAGCGCGAGCGCGACATG-dt(30)BN-3' (where B = A,
         C, or G and N = A, C, G, or T). Average insert size 1.9
         kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
         by PCR. This library was enriched for full-length clones
         and was constructed by Clontech Laboratories (Palo Alto,
         CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.3%; Score 21; DB 12; Length 557;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATGCGTCTG 21
|||||
Db 256 CCTGTGGTCTATGCGTCTG 236
|||||

```

```

RESULT 37
CD687141/c
LOCUS          CD687141          557 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION    EST3662 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION     CD687141
VERSION       CD687141.1  GI:32204715
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 557)
AUTHORS       Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Fan,Z.-G. and
               Zeng,Y.-X.
TITLE         Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL       Unpublished (2003)
COMMENT       Contact: YiXin Zeng
               Cancer Center
               Sun Yat-sen University
               651 Dongfeng Road East, Guangzhou 510060, China
               Tel: 86-1380-9770-743
               Fax: 86-20-8775-4506
               Email: yxzeng@gzsums.edu.cn.
FEATURES      Location/Qualifiers
               source          1..557
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /tissue_type="normal nasopharynx"
               /clone_lib="human nasopharynx"
               /note="ESTs generated from a normal nasopharynx cDNA
               library from southern Chinese"
ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 557;
Best Local Similarity 100.0%; Pred.No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CCCTGTTGGTCTATGCGTCTG  21
      |||||||
DB   160 CCCTGTTGGTCTATGCGTCTG  140

RESULT 38
CD700948/c
LOCUS          CD700948          559 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION    EST17504 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION     CD700948
VERSION       CD700948.1  GI:32231610
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 559)
AUTHORS       Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Fan,Z.-G. and
               Zeng,Y.-X.
TITLE         Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL       Unpublished (2003)
COMMENT       Contact: YiXin Zeng
               Cancer Center
               Sun Yat-sen University
               651 Dongfeng Road East, Guangzhou 510060, China
               Tel: 86-1380-9770-743
               Fax: 86-20-8775-4506
               Email: yxzeng@gzsums.edu.cn.
FEATURES      Location/Qualifiers
               source          1..559
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /tissue_type="normal nasopharynx"

```

```

/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"
ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 559;
Best Local Similarity 100.0%; Pred.No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CCCTGTTGGTCTATGCGTCTG  21
      |||||||
DB   274 CCCTGTTGGTCTATGCGTCTG  254

RESULT 39
CB265399/c
LOCUS          CB265399          563 bp      mRNA      linear      EST 20-FEB-2003
DEFINITION    1004304 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
               cDNA 5', mRNA sequence.
ACCESSION     CB265399
VERSION       CB265399.1  GI:28439987
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 563)
AUTHORS       Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE         EST analysis of human adipose gene expression
JOURNAL       Unpublished (2002)
COMMENT       Contact: Gong Da-Wei
               Division of Endocrinology, Diabetes and Nutrition
               University of Maryland
               660 Redwood St, HR497, Baltimore, MD 21201, USA
               Tel: 410 706 1672
               Fax: 410 706 1622
               Email: dgong@medicine.umaryland.edu
               PCR Primers
               FORWARD: CTCGGGAGCGCGCATTTGTTGGT
               BACKWARD: AATACGACTCATTATAGGCGCAATGG
               Seq primer: GTTGTACCCGGGAATTC.
               Location/Qualifiers
               source          1..563
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /sex="Male and Female"
               /tissue_type="Adipose"
               /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
               /note="Vector: lambdaTriplex"
ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 563;
Best Local Similarity 100.0%; Pred.No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CCCTGTTGGTCTATGCGTCTG  21
      |||||||
DB   214 CCCTGTTGGTCTATGCGTCTG  194

RESULT 40
CA942442/c
LOCUS          CA942442          571 bp      mRNA      linear      EST 30-DEC-2002
DEFINITION    ir58a06.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6549227 5'
               similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY
               ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.
ACCESSION     CA942442
VERSION       CA942442.1  GI:27430922
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia: Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 571)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Rorko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.

TITLE  
JOURNAL  
COMMENT  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other ESTs: 158a06.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@oicp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@wustl.edu)

Seq primer: -40UP from Gibco  
High quality sequence stop: 487.

FEATURES  
source  
1..571  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6549227"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:  
NotI; Site 2: XhoI; cDNA made by oligo-dr priming.  
Size selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 571;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATCGCTCTG 21  
|||||  
DB 214 CCCTGTTGGTCTATCGCTCTG 194

RESULT 41  
EW831052/c  
LOCUS  
DEFINITION  
K-BST0104933 S14K402s1 Homo sapiens cDNA clone S14K402s1-18-C12 5',  
mRNA sequence.

ACCESSION  
EW831052  
VERSION  
EW831052.1 GI:19187461  
KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 579)  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Kim, Y.S.

TITLE  
JOURNAL  
COMMENT  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-gong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 18 row: C column: 12  
High quality sequence stop: 579.  
Location/Qualifiers  
1..579  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S14K402s1-18-C12"  
/cell\_line="K402"  
/lab\_host="Top10F"  
/clone\_lib="S14K402s1"  
/note="Organ: Stomach; Vector: pT18RP1; Site 1: EcoRI;  
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library. After analyzing and  
sequencing about 2,000 ~ 3,000 colonies in original cDNA  
library, the abundant cDNAs were selected and amplified by  
PCR reaction using vector region primer including T7  
promoter as 5' primer and N(dT)14 as 3' primer. The PCR  
products were used as template for synthesis of  
biotinylated single stranded RNA by in vitro transcription  
reaction. The synthesized RNA probes were hybridized with  
antisense single stranded cDNAs prepared from original  
library and incubated with avidin-gel. After removing  
DNA-RNA hybrids by centrifuge, the subtracted cDNA  
libraries were constructed by transfection of the  
remaining DNA into competent cells E. coli Top10F with  
electroporation method."

FEATURES  
source

ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 579;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATCGCTCTG 21  
|||||  
DB 170 CCCTGTTGGTCTATCGCTCTG 150

RESULT 42  
EW737984/c

LOCUS  
DEFINITION  
K-EST0002067 S1SNUS Homo sapiens cDNA clone S1SNUS-25-B01 5', mRNA  
sequence.

ACCESSION  
EW737984  
VERSION  
EW737984.1 GI:19059313  
KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
1 (bases 1 to 582)  
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.



TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 25 row: B column: 01  
High quality sequence stop: 582.

## FEATURES

Location/Qualifiers  
1. .582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S1SNUS-25-B01"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="lymphoblast-like"  
/cell\_line="SNU-5"  
/lab\_host="Top10F"  
/clone\_lib="S1SNUS"  
/note="Organ: Stomach; Vector: pCNS; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

source

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 582;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
|||||

Db 284 CCTGTGTCATGCGTCTG 264  
|||||

## RESULT 44

BU783392/c

## LOCUS

DEFINITION

BU783392

ACCESSION

VERSION

SOURCE

ORGANISM

REFERENCE

AUTHORS

KEYWORDS

COMMENT

Query Match 100.0%; Score 21; DB 12; Length 582;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
|||||

Db 233 CCTGTGTCATGCGTCTG 213  
|||||

## RESULT 43

AV706521/c

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AV706521 585 bp mRNA linear EST 09-OCT-2000  
AV706521 ADB Homo sapiens cDNA clone ADBAVE11 5', mRNA sequence.  
AV706521  
EST.  
AV706521.1 GI:107233800

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 585)  
Peng, Y., Song, H., Huang, Q., Gu, Y., Yang, Y., Gao, G.,  
Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z.,  
Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,  
Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.

Homo sapiens cDNA ADB clones  
Unpublished (2000)  
Contact: Zengqiang Han

Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)

Fax: 86-21-50801922

Email: hanzg@hgc.sh.cn

This clone is available at CHGC in Shanghai.

## FEATURES

source

Location/Qualifiers  
1. .585  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ADBAVE11"  
/tissue\_type="Adrenal gland"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="ADBS"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
|||||

Db 284 CCTGTGTCATGCGTCTG 264  
|||||

## RESULT 44

BU783392/c

## LOCUS

DEFINITION

BU783392

ACCESSION

VERSION

SOURCE

ORGANISM

REFERENCE

AUTHORS

KEYWORDS

COMMENT

Query Match 100.0%; Score 21; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
|||||

Db 284 CCTGTGTCATGCGTCTG 264  
|||||

## RESULT 43

AV706521/c

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AV706521 585 bp mRNA linear EST 11-OCT-2000  
AV706521 ADB Homo sapiens cDNA clone IMAGE:6123426  
AV706521  
EST.  
AV706521.1 GI:23827516

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 585)  
Melton, D., Brown, J., Kent, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Konko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishevili, R.,  
Williams, T., Jackson, Y. and Bowers, Y.

Homo sapiens cDNA ADB clones  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: 40RP from Gibco  
High quality sequence stop: 429.  
Location/Qualifiers  
1. .585  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6123426"

/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:

XhoI; Site 2: EcoRI; Constructed with Lambda ZapII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol

(http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 100.0%; Score 21; DB 13; Length 585;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
DB 167 CCTGTGGTCTATGCGTCTG 147

## RESULT 45

BI911442/c

LOCUS

DEFINITION

603063356F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212669 5',  
mRNA sequence.

ACCESSION

BI911442

VERSION

BI911442.1

KEYWORDS

357

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 585)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium, Inc.

C-one distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLML1533 row: p column: 14

High quality sequence start: 18

High quality sequence stop: 585.

Location/Qualifiers

1..588

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5212669"

/tissue\_type="leukocyte"

/lab\_host="DH10B"

/clone\_lib="NIH MGC 118"

/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of

non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 027. Note:

this is a NIH\_MGC Library."

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 24;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 24;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21

|||||

DB 208 CCTGTGGTCTATGCGTCTG 188

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
DB 225 CCTGTGGTCTATGCGTCTG 205

## RESULT 46

BM876262/c

LOCUS

DEFINITION

BM876262  
3' similar to SW:HA2Q HUMAN P20036 H1A CLASS II HISTOCOMPATIBILITY

ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

ACCESSION

BM876262

VERSION

BM876262.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 592)

AUTHORS

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blais, A.,

Schmitt, A., Treising, B., Ritter, E., Ronko, T., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagazeishvili, R.,

Williams, T., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other\_Estis: l160C02.y1

Contact: Douglas Melton, Klaus H. Kaestner, &amp; Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bichp.harvard.edu

Library was constructed by Dr. J. Ferrer in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center for information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco

High quality sequence stop: 482.

Location/Qualifiers

1..592

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5633202"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:

XhoI; Site\_2: EcoRI; Constructed with lambda ZapII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

(http://genome.wustl.edu/est/lambda\_protocol.shtml).

Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt

Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this

is a Washington University Pancreas EST project library."

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 24;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 47
BG926106/c
LOCUS
DEFINITION
HNC50-1-B12.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
sequence.
ACCESSION
BG926106
VERSION
BG926106.1 GI:14320629
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 594)
AUTHORS
Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
TITLE
Identification and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteochondritic cartilage cDNA libraries
JOURNAL
Osteoarthr. Cartil. 9 (7), 641-653 (2001)
MEDLINE
21482651
PubMed
11597177
COMMENT
Contact: Sanjay Kumar
UW2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@gsk.com
Seq primer: T7.
FEATURES
source
1..594
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/clone_lib="HNC (Human Normal Cartilage)"
/note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
Directional"
ORIGIN
Query Match 100.0%; Score 21; DB 12; Length 594;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 366 CCCTGTTGGTCTATGCGTCTG 346

RESULT 48
CD693703/c
LOCUS
DEFINITION
EST102226 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION
CD693703
VERSION
CD693703.1 GI:32217621
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 595)
AUTHORS
Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE
Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL
Unpublished (2003)
COMMENT
Contact: YiXin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-1380-9770-743
Email: yxzeng@gzsums.edu.cn.
FEATURES
source
1..597
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 597;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCGTGTTGGTCTATGCGTCTG 21
|||||
Db 227 CCCGTGTTGGTCTATGCGTCTG 207

RESULT 50
CD695219/c
LOCUS
DEFINITION
EST11742 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION
CD695219
VERSION
CD695219.1 GI:32220597
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

```

Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

```

## ORIGIN

```

Query Match 100.0%; Score 21; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 238 CCCTGTTGGTCTATGCGTCTG 218

```

```

RESULT 49
CD684394/c
LOCUS
DEFINITION
EST914 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION
CD684394
VERSION
CD684394.1 GI:32199345
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 597)
AUTHORS
Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE
Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL
Unpublished (2003)
COMMENT
Contact: YiXin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
FEATURES
source
1..597
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 597;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCCGTGTTGGTCTATGCGTCTG 21
|||||
Db 227 CCCGTGTTGGTCTATGCGTCTG 207

RESULT 50
CD695219/c
LOCUS
DEFINITION
EST11742 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION
CD695219
VERSION
CD695219.1 GI:32220597
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
Zeng, Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixir Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.

FEATURES
source 1..598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 598;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATGCGTCG 21
|||||
Db 256 CCCTGTTGGTCTATGCGTCG 236

RESULT 51
BQ270752/c
LOCUS BQ270752 599 bp mRNA linear EST 15-JUL-2003
DEFINITION iX0401.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:5779993
S, similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY
ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.
ACCESSION BQ270752
VERSION BQ270752.1 GI:20495818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 599)
AUTHORS Yelton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R.,
Williams, T., Jackson, Y. and Bowers, Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ik0401.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1852
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer in vivo mass-excised to
pBluescript SK- by Dr. H Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 430.
Location/Qualifiers

FEATURES
source 1..598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

```

```

source 1..599
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5779993"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: i:
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. E. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN
Query Match 100.0%; Score 21; DB 13; Length 599;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATGCGTCG 21
|||||
Db 167 CCCTGTTGGTCTATGCGTCG 147

RESULT 52
BQ725429/c
LOCUS BQ725429 601 bp mRNA linear EST 05-JAN-2001
DEFINITION bx15h06.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo
sapiens cDNA clone bx15h06 5', mRNA sequence.
ACCESSION BQ725429
VERSION BQ725429.1 GI:12041348
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 601)
AUTHORS Wistow, G. J., Bernstein, S., Behai, A. and Smith, D.
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics.
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 456 0078
Email: graeme@helix.nih.gov
Plate: 15 row: h column: 06
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bx15h06"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Un-normalized, unamplified):
BX"
/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris
tissue was pooled from 10 individuals ranging in age from
4-80 years and RNA was extracted. From this pooled sample
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A
directionally cloned cDNA library in the pCMVSPORT6 vector
was constructed at Life Technologies, essentially
following the protocols of the SuperScript plasmid System

```

FEATURES

full details of which are contained in the manufacturer's Instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor

[5'-pGACTAGTTCTAGATCGGAGCGCGCC(T)15-3']. Not I blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 601;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 157 CCTGTGGTCTATGCGTCTG 137

## RESULT 53

BU783579/c  
LOCUS  
DEFINITION in06c10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123643  
5' similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
ANTIGEN, DP ALPEA CHAIN PRECURSOR ; mRNA sequence.

ACCESSION BU783579.1 GI:23827903

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 602)

Melton, D., Brown, G., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Demishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., R.,

Williams, T., Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other ESTs: in06c10.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -4ORP from Gibco

High quality sequence stop: 431.

Location/Qualifiers

1..602

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6123643"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/notes="Organ: pancreas; Vector: pBluescript SK-; Site 1:

XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

([http://genome.wustl.edu/est/lambda\\_protocol.shtml](http://genome.wustl.edu/est/lambda_protocol.shtml)).

Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt

## ORIGIN

Query Match 100.0%; Score 21; DB 13; Length 602;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 203 CCTGTGGTCTATGCGTCTG 183

## RESULT 54

CA405960/c  
LOCUS  
DEFINITION 1002109 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
cDNA 5', mRNA sequence.

ACCESSION CA405960.1 GI:24770831

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 610)

Yang, R.-Z., Shuldiner, A. and Gong, D.-W.

EST analysis of human adipose gene expression

Unpublished (2002)

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, HH497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGAGCGCGCCATGTGTGTGGT

BACKWARD: AATACGACTCATATAGCGGATTTGG

Seq primer: GTTGTACCCGGGAATTC.

Location/Qualifiers

1..610

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/sex="Male and Female"

/tissue\_type="Adipose"

/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

/note="Vector: lambdaTriplex"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 610;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 203 CCTGTGGTCTATGCGTCTG 183

## RESULT 55

CD686880/c  
LOCUS  
DEFINITION EST3401 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD686880

VERSION CD686880.1 GI:32204209

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 618)

**AUTHORS** Liu, X.-Q., Zhot, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.  
**TITLE** Transcriptional Gene Expression Profile of Human Nasopharynx  
**JOURNAL** Unpublished (2003)  
**COMMENT** Contact: Yixin Zeng  
 Cancer Center  
 Sun Yat-sen University  
 651 Dongfeng Road East, Guangzhou 510060, China  
 Tel: 86-1380-9770-743  
 Fax: 86-20-8775-4506  
 Email: yxzeng@gzsums.edu.cn.

**FEATURES**  
 source  
 1..618  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 283 CCTGTGGTCTATGCGTCTG 263

**RESULT 56**  
**LOCUS** BF128930/c  
**DEFINITION** 601811229F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4054350 5', mRNA sequence.

**ACCESSION** BF128930  
**VERSION** BF128930.1 GI:10967970  
**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 623)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov

**REFERENCE**  
**AUTHORS** Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
**TITLE** cDNA Library Preparation: Ling Hong/Rubin Laboratory  
**JOURNAL** cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
**COMMENT** DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

**FEATURES**  
 source  
 1..623  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4054350"  
 /tissue\_type="primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_48"  
 /note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 623;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 55 CCTGTGGTCTATGCGTCTG 35

**RESULT 57**  
**LOCUS** BG545380/c

**DEFINITION** 60257261F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4700906 5', mRNA sequence.

**ACCESSION** BG545380

**VERSION** BG545380.1 GI:13543873

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 626)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov

**REFERENCE** Tissue Procurement: CLONTECH Laboratories, Inc.

**AUTHORS** cDNA Library Preparation: CLONTECH Laboratories, Inc.

**TITLE** cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

**JOURNAL** DNA Sequencing by: Incyte Genomics, Inc.

**COMMENT** Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

Plate: LCM1535 row: e column: 03  
 High quality sequence stop: 626.

**FEATURES**  
 source  
 1..626  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4700906"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcctggcc); Site 2: SfiI (ggcgcattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCCGCCACATG-dt(30)EN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 21; DB 12; Length 626;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
 |||||  
 Db 220 CCTGTGGTCTATGCGTCTG 200

**RESULT 58**

**LOCUS** BG715375/c

627 bp mRNA linear EST 08-MAY-2001

```

DEFINITION 602675513F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4798153 5',
mRNA sequence.
ACCESSION BG715375
VERSION BG715375.1 GI:13994558
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
REFERENCE 1 (bases 1 to 627)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cchapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIA10685 row: i column: C4
High quality sequence stop: 627.
Location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798155"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH MGC 96"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 21; DB 12; Length 627;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTGCTATGCGTCTG 21
|||||
DB 215 CCTGTGTGCTATGCGTCTG 195
|||||

RESULT 59
CB267128/c
LOCUS 627 bp mRNA linear EST 20-FEB-2003
DEFINITION 1006034 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
cDNA 5', mRNA sequence.
ACCESSION CB267128
VERSION CB267128.1 GI:28441714
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 627)
AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
JOURNAL EST analysis of human adipose gene expression
Unpublished (2002)
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition

```

```

University of Maryland
660 Redwood St., HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGCCCATCTGTGTGGT
BACKWARD: AATACGACTACTATAGGCGGAATTGG
Seq primer: GTTGGTACCGGAATTC.
Location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriplex"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 627;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTGTGCTATGCGTCTG 21
|||||
DB 158 CCCTGTGTGCTATGCGTCTG 138
|||||

RESULT 60
CD691540/c
LOCUS 629 bp mRNA linear EST 25-JUN-2003
DEFINITION EST8063 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD691540
VERSION CD691540.1 GI:32213359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 629)
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
Zeng, Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 DongFeng Road East, Guangzhou 510060, China
Tel: 86-1390-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@zsus.edu.cn.
Location/Qualifiers
1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 629;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTGTGCTATGCGTCTG 21
|||||
DB 254 CCCTGTGTGCTATGCGTCTG 234
|||||

```

Search completed: April 20, 2004, 12:33:44

Job time : 1337.59 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Ran on: April 20, 2004, 09:33:59 ; Search time 2387 Seconds  
(without alignments)  
7937.369 Million cell updates/sec

Title: US-09-877-819B-55

Perfect score: 254

Sequence: 1 atcaaggcggaccatgtgtc.....acactcaggccaccaccgat 254

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.red.\*

36: em.htg.man.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	241.4	95.0	466	6	BD058339 Secreted
2	241.4	95.0	476	6	AX884252 Sequence
3	241.4	95.0	476	6	BD023862 Sequence
4	241.4	95.0	576	6	AX884251 Sequence
5	241.4	95.0	576	6	BD023861 Sequence
6	241.4	95.0	661	6	E00485 DNA sequence
7	241.4	95.0	661	6	I03086 Sequence 5
8	241.4	95.0	690	9	S40633 HLA class I
9	241.4	95.0	818	9	M27487 Homo sapien
10	241.4	95.0	1048	9	X00457 Human mRNA
11	241.4	95.0	1140	6	I03088 Sequence 7
12	241.4	95.0	1201	6	I03006 Sequence 6
13	241.4	95.0	1201	6	I03423 Sequence 6
14	241.4	95.0	1259	6	AX552229 Sequence
15	239.8	94.4	1201	6	E00484
16	233.8	92.0	2986	9	HUMHRC02
17	233.8	92.0	14646	9	HSKLABSA
18	233.8	92.0	64360	9	EX120009 Human DNA
19	233.8	92.0	106728	9	AL805913 Human DNA
20	233.8	92.0	124899	9	AL645931 Human DNA
21	233.8	92.0	181228	2	AC011086
22	233.8	92.0	187964	9	AL662824 Human DNA
23	233.2	91.8	252	9	HS248473
24	232.4	91.5	246	9	HS0103X2
25	230.8	90.9	246	9	HSKLABSA
26	230.2	90.6	257	9	AF165160
27	229.2	90.2	246	9	AF346471
28	229.2	90.2	326	9	AF013767
29	228.6	90.0	1201	9	BC009956 Homo sapi
30	227.4	89.5	286	9	HS087556
31	226.4	89.1	1480	6	AX780118
32	226.4	89.1	1480	6	AX780119
33	226.2	89.1	279	9	AF076284
34	225.8	88.9	287	9	AF015395
35	224.4	88.3	396	6	BD058253 Secreted
36	222.6	87.6	268	9	HSKLABSA
37	222.1	87.0	244	9	AF098794
38	219.8	86.5	246	9	HS0201X2
39	219.8	86.5	279	9	AF076285
40	219.2	86.3	258	9	AF118120
41	214	84.3	265	9	AF092049
42	209.6	82.5	249	9	HSKLABSA
43	206.8	81.4	222	9	AF074848
44	206	81.1	226	9	HSKLABSA
45	202.8	79.8	214	9	HUMHRC11A
46	199.6	78.6	220	9	HUMHRC11A
47	198.2	78.0	226	9	HSKLABSA
48	196.6	77.4	226	9	HSKLABSA
49	195.2	76.9	222	9	AF074847
50	195	76.8	226	9	HSKLABSA
51	191.6	75.4	198	9	HUMHRC11A
52	189	74.4	189	9	HSKLABSA
53	186.8	73.5	198	9	HSKLABSA
54	185.2	72.9	198	9	HSKLABSA
55	184.2	72.5	189	9	AF026694
56	184	72.4	192	9	HUMHRC11A
57	183.6	72.3	198	9	HUMHRC11A
58	181	71.3	189	9	AF026707
59	177.8	70.0	189	9	AF026693
60	176.2	69.4	189	9	AF026705
61	174.6	68.7	189	9	AF026701
62	174.6	68.7	189	9	AF026702
63	174.6	68.7	189	9	AF026703
64	174.6	68.7	189	9	AF026706
65	173	68.1	189	9	AF026695

66	173	68.1	189	9	AF026697	Pongo pyg
67	173	68.1	189	9	AF026704	Macaca fa
68	171.4	67.5	189	9	AF026696	AF026704 Pongo pyg
69	169.4	66.7	1263	10	AB084158	AB084158 Cavia por
70	166.6	65.6	189	9	AF529200	AF529200 Aotus ran
71	165.4	65.1	349980	6	AX344564	AX344564 Sequence
72	163.4	64.3	189	9	AF026699	AF026699 Saimiri s
73	161.8	63.7	189	9	AF026700	AF026700 Saimiri s
74	160.2	63.1	189	9	AF026698	AF026698 Saimiri s
75	158	62.2	192	9	HUMHEC11TF	L31624 Human MHC c
76	157.2	61.9	267	6	BD058170	BD058170 Secreted
77	146	57.5	267	6	AX237167	AX237167 Sequence
78	146	57.5	267	6	AX237352	AX237352 Sequence
79	146	57.5	272	6	AX237066	AX237066 Sequence
80	146	57.5	294	6	AX237304	AX237304 Sequence
81	146	57.5	294	6	AX237554	AX237554 Sequence
82	144.8	57.0	349980	6	AX344560	AX344560 Sequence
83	141	55.5	2763	4	RABMDHPA	M22640 Rabbit MHC
84	139	54.7	87684	4	AY152828	AY152828 Felis cat
85	139	54.7	132445	4	AY152826	AY152826 Felis cat
86	137.4	54.1	1834	9	HSSBA2P	X01634 Human HLA-S
87	137.4	54.1	22847	9	HSD0826Q1	AL049813 Human DNA
88	135.8	53.5	93842	9	AL845446	AL845446 Human DNA
89	129	50.8	6565	9	HUMHE5XA	M11591 Human MHC c
90	116.6	45.9	924	4	FCU51578	U15178 Felis catus
91	115.6	45.5	744	9	HUMHEDRAN	M60333 Human MHC c
92	115.6	45.5	747	6	AR016516	AR016516 Sequence
93	115.6	45.5	747	6	AR096899	AR096899 Sequence
94	115.6	45.5	747	6	BD136094	BD136094 Vaccines
95	115.6	45.5	750	6	BD137956	BD137956 Monoclonal
96	115.6	45.5	819	9	HUMHEDR3M	M60134 Human MHC c
97	115.6	45.5	1117	10	SOIMHCDFAF	M97624 Sciurus abe
98	115.6	45.5	1183	6	AX303522	AX303522 Sequence
99	115.6	45.5	1183	9	HUMHEDRAM	K01171 Human HLA-D
100	115.6	45.5	1446	6	BD137962	BD137962 Monoclonal
101	115.6	45.5	1851	6	BD137963	BD137963 Monoclonal
102	114	44.9	369	6	AX351328	AX351328 Sequence
103	114	44.9	491	6	BD071896	BD071896 Secreted
104	114	44.9	491	6	AX887895	AX887895 Sequence
105	114	44.9	509	6	AX887895	AX887895 Sequence
106	114	44.9	509	6	BD027505	BD027505 Sequence
107	114	44.9	542	6	AX884324	AX884324 Sequence
108	114	44.9	542	6	BD023934	BD023934 Sequence
109	114	44.9	1099	6	I03005	I03005 Sequence 5
110	114	44.9	1099	6	I03054	I03054 Sequence 1
111	114	44.9	1099	6	I03087	I03087 Sequence 6
112	114	44.9	1099	6	I03414	I03414 Sequence 5
113	114	44.9	1099	6	I03715	I03715 Sequence 6
114	114	44.9	1199	9	HUMHEDR4	J00194 human hla-d
115	114	44.9	1244	6	AR096965	AR096965 Sequence
116	114	44.9	1244	6	AR096966	AR096966 Sequence
117	114	44.9	1244	6	AR096967	AR096967 Sequence
118	114	44.9	1244	6	AR096968	AR096968 Sequence
119	114	44.9	1244	6	AR367302	AR367302 Sequence
120	114	44.9	1244	6	AR367303	AR367303 Sequence
121	114	44.9	1244	6	AR367304	AR367304 Sequence
122	114	44.9	1244	6	AR367305	AR367305 Sequence
123	114	44.9	1289	9	BC032350	BC032350 Homo sapi
124	113.2	44.6	1121	4	FCU51576	U51576 Felis catus
125	113.2	44.6	1121	4	SOIMHCDDRAA	M97619 Sciurus abe
126	113.2	44.6	1155	10	SOIMHCDDRA	M97620 Sciurus abe
127	113.2	44.6	1163	10	SOIMHCDDRA	M97621 Sciurus abe
128	113.2	44.6	1163	10	SOIMHCDDRA	M97623 Sciurus abe
129	113.2	44.6	1163	10	SOIMHCDDRA	M97625 Sciurus abe
130	113.2	44.6	1163	10	SOIMHCDDRA	M97626 Sciurus abe
131	113.2	44.6	1163	10	SOIMHCDDRA	M97628 Sciurus abe
132	113.2	44.6	1164	10	SOIMHCDDRA	M97622 Sciurus abe
133	113.2	44.6	1164	10	SOIMHCDDRA	M97627 Sciurus abe
134	113.2	44.6	1166	10	SOIMHCDDRA	M97629 Sciurus abe
135	113.2	44.6	1179	10	SOIDRA	M97615 Sciurus abe
136	112.2	44.2	130898	4	AY152836	AY152836 Felis cat
137	112	44.1	292	10	S8039CS2	S80392 MHC RT1.H a
138	111.6	44.0	762	4	AB008754	AB008754 Capra hir
139	111.2	43.8	148248	2	RN235J17	AL603722 Rattus no
140	111.2	43.8	251251	2	AC120734	AC120734 Rattus no
141	110.6	43.5	74680	4	AY152834	AY152834 Felis cat
142	110.6	43.5	77680	4	AY152835	AY152835 Felis cat
143	110.2	43.4	246	4	HRSDRAC	L47172 Equus caball
144	110.2	43.4	246	4	DR201322	D42013 Rattus norv
145	110.2	43.4	246	4	AY285936	AY285936 Sus scrof
146	110.2	43.4	759	4	AY285939	AY285939 Sus scrof
147	110.2	43.4	805	4	AY247783	AY247783 Sus scrof
148	110.2	43.4	807	4	AY191779	AY191779 Sus scrof
149	110.2	43.4	821	9	MACDRA	L27739 Macaca mula
150	110.2	43.4	1205	4	SHPDRA	J07983 Ovis aries
151	109.8	43.2	1806	9	HUMHEDRH	J00201 human hla-d
152	109.8	43.2	3742	6	HUMHEDRS02	C00204 Human HLA-D
153	109.8	43.2	5724	6	AX337215	AX337215 Sequence
154	109.8	43.2	5724	6	HS107	X00274 Human gene
155	109.8	43.2	95263	9	AL935032	AL935032 Human DNA
156	109.8	43.2	109682	9	AX120007	AX120007 Human DNA
157	109.8	43.2	131012	9	AL662796	AL662796 Human DNA
158	109.8	43.2	131234	9	HS172K2	Z84844 Human DNA s
159	109.8	43.2	150623	9	AL670296	AL670296 Human DNA
160	109.8	42.9	1962	4	CAOLADRA2	Z11600 O.aries gen
161	109.8	42.9	3519	4	AB008755	AB008755 Capra hir
162	108.6	42.8	246	4	HRSDRAB	L47171 Equus asinu
163	108.6	42.8	246	4	HRSDRAB	L47173 Equus hemio
164	108.6	42.8	246	4	AY285928	AY285928 Sus scrof
165	108.6	42.8	754	4	AY243102	AY243102 Sus scrof
166	108.6	42.8	759	4	AY243106	AY243106 Sus scrof
167	108.6	42.8	759	4	AY285929	AY285929 Sus scrof
168	108.6	42.8	759	4	AY285937	AY285937 Sus scrof
169	108.6	42.8	759	4	AY285938	AY285938 Sus scrof
170	108.6	42.8	759	4	AY459302	AY459302 Sus scrof
171	108.6	42.8	765	4	AY285926	AY285926 Sus scrof
172	108.6	42.8	765	4	AY285933	AY285933 Sus scrof
173	108.6	42.8	765	4	AY285935	AY285935 Sus scrof
174	108.6	42.8	805	4	AY247781	AY247781 Sus scrof
175	108.6	42.8	805	4	AY247782	AY247782 Sus scrof
176	108.6	42.8	1051	4	BOVK01A	D37955 Bovine BoLA
177	108.6	42.8	1195	4	BOVHECII	X78308 B.taurus (F
178	108.6	42.8	1222	4	BOVHECII	D37956 Bovine BoLA
179	108.6	42.8	1256	4	PIGMHCIIID	M93028 Pig MHC cla
180	108.6	42.8	1300	4	PIGMHCIIID	M92445 Pig MHC cla

## ALIGNMENTS

RESULT 1	BD058339	466 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD058339	Secreded	expressed	sequence tags (seSTs).	
DEFINITION	BD058339	Secreded	expressed	sequence tags (seSTs).	
ACCESSION	BD058339	GI:22603945			
VERSION	BD058339.1	GI:22603945			
KEYWORDS	JP 2001519666-A/194.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	1 (bases 1 to 466)				
AUTHORS	Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D., Treacy,M., Spaulding,V. and Agostino,M.J.				
TITLE	Secreded expressed sequence tags (seSTs).				
JOURNAL	Patent: JP 2001519666-A 194 23-OCT-2001;				
COMMENT	GENETICS INSTITUTE INC				
	PN JP 2001519666-A/194				
	PD 23-OCT-2001				
	PF 10-APR-1998 JP 1998543068				
	PR 10-APR-1997 US 08/835913				
	PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI DAVID MERBERG,				
	PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC				
	C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC				
	Strandedness:				

```

Double;
CC Topology: Linear; Location/Qualifiers.
FH Key Location/Qualifiers
FEATURES
    source
        1..466
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /db_xref="taxon:4577"
ORIGIN
    Query Match          95.0%; Score 241.4; DB 6; Length 466;
    Best Local Similarity 99.2%; Pred. No. 3.8e-63;
    Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 60
Db 113 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 172
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 120
Db 173 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 232
QY 121 ACCGTCGTCATCTGAGGAGTTTGCCCAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 180
Db 233 ACCGTCGTCATCTGAGGAGTTTGCCCAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 292
QY 181 GCTAACATTGCTATATGAACAACAACCTTGATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 293 GCTAACATTGCTATATGAACAACAACCTTGATACCTTGATCCAGCGTTCC-ACCACACT 352
QY 240 CAGGCCACCAACGAT 254
Db 353 CAGGCCACCAACGAT 367

RESULT 2
AX884252
LOCUS AX884252 476 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 115 from Patent EP1033401.
ACCESSION AX884252
VERSION AX884252.1 GI:40039227
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
    1 Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
    Expressed sequence tags and encoded human proteins
    Patent: EP 1033401-A 115 06-SEP-2000;
    Gense: (FR)
FEATURES
    source
        1..476
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAE998356.1"
            /db_xref="GI:40039228"
            /translation="MRPEDEMFHILRAVILSLAFLLSLRGAGAIKADHVSVAAPV
            QTRPFGTFPEFDEDMFYVDLKKETVHLSFEGQAPSFBAQGGLAN-ALLNNLN
            TLIORSNHTQATNDPPETVFPKPEP"
            95..187
            /note="score 10.1 seq SLAFLLSLRGAGA/IK"
    sig_peptide
        95..187
ORIGIN
    Query Match          95.0%; Score 241.4; DB 6; Length 476;
    Best Local Similarity 99.2%; Pred. No. 3.8e-63;
    Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 60
Db 113 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 172
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 120
Db 173 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 232
QY 121 ACCGTCGTCATCTGAGGAGTTTGCCCAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 180
Db 233 ACCGTCGTCATCTGAGGAGTTTGCCCAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 292
QY 181 GCTAACATTGCTATATGAACAACAACCTTGATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 293 GCTAACATTGCTATATGAACAACAACCTTGATACCTTGATCCAGCGTTCC-ACCACACT 352
QY 240 CAGGCCACCAACGAT 254
Db 353 CAGGCCACCAACGAT 367

RESULT 3
BD023862
LOCUS BD023862 476 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD023862
VERSION BD023862.1 GI:22565085
KEYWORDS JP 2001269182-A/108.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
    1 (bases 1 to 476)
    Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    Patent: JP 2001269182-A 108 02-OCT-2001;
    GENSET
    OS Homo sapiens (human)
    PN JP 2001269182-A/108
    PD 02-OCT-2001
    PF 24-FEB-2000 JP 2000118773
    PR 26-FEB-1999 US 60/122487
    PT JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
    PI JORDAN
    PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
    C12N5/10.
    PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
    G06F15/40
    CC score 10.1
    CC seq SLAFLLSLRGAGA/IK
    FH Key Location/Qualifiers
    FT CDS
    PT sig_peptide 95..187.
    FT Location/Qualifiers
        1..476
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
    Query Match          95.0%; Score 241.4; DB 6; Length 476;
    Best Local Similarity 99.2%; Pred. No. 3.8e-63;
    Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 60
Db 188 ATCAAGCGGACCATGTCTCAACTATATGCGCGTTTGACAGCATAGACCAACAGGG 247
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 120
Db 248 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTGGAACAAGAGGAG 307
QY 121 ACCGTCGTCATCTGAGGAGTTTGCCCAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 180

```

```

Db 308 ACCGCTGGGCACTGGAGGAGTTGGCCACAGCCTTTCTCTTTCAGGCTCAGGGCGGGCTG 367
Qy 181 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCC-ACCACACT 239
|||||
Db 368 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCCACCACT 427
Qy 240 CAGGCCACCAACCGAT 254
Db 428 CAGGCCACCAACCGAT 442

RESULT 4
AX884251
LOCUS AX884251 576 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 114 from Patent EP1033401.
ACCESSION AX884251
VERSION AX884251.1 GI:40039225
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Dumas Milne Edwards,J.B., Ducloux,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 114 06-SEP-2000;
Genset (FR)
FEATURES
source 1..576
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
195..>575
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAE98365.1"
/db_xref="GI:40039226"
/translation="MRPEDMEHTRAVILRALSLAFLLSLRGAGA,KADHVSTYAFV
QTHRTPTGEFPEDEDEMFYVDLKKTVHLSFQGFSEAGQGLLANTAILNNLN
TLIORSNHTQATNDPPETVFPKPEP"
sig_peptide 195..287
/notes="score 10.1 seq SLAFLLSLRGAGA/IK"

ORIGIN
Query Match 95.0%; Score 241.4; DB 6; Length 576;
Best Local Similarity 99.2%; Pred. No. 3.9e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGGACCACTGTGCAACTTATGCGCGTTTGACAGCGCATAGACCAACAGGG 60
Db 288 ATCAAGCGGACCACTGTGCAACTTATGCGCGTTTGACAGCGCATAGACCAACAGGG 347

Qy 61 GAGTTTATGTTGAATTGTGAGAGATGAGATGTTCTATCTGATCTGCAAGAAGAGG 120
Db 348 GAGTTTATGTTGAATTGTGAGAGATGAGATGTTCTATCTGATCTGCAAGAAGAGG 407

Qy 121 ACCGCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 180
Db 408 ACCGCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 467

Qy 181 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCC-ACCACACT 239
Db 468 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCCACCACT 527

Qy 240 CAGGCCACCAACCGAT 254
Db 528 CAGGCCACCAACCGAT 542

RESULT 5
BD023861
LOCUS BD023861 576 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.

```

```

ACCESSION BD023861
VERSION BD023861.1 GI:22565084
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Edwards,J.B.D.M., Ducloux,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 107 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN CP 2001269182-A/107
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
P2 JORDAN
PC C12N15/09,C07K14/435,C07K16/19,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC score 10.1
CC seq SLAFLLSLRGAGA/IK
PH key location/Qualifiers
FT CDS 195..575
FT sig_peptide 195..287.
FEATURES
source 1..576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 95.0%; Score 241.4; DB 6; Length 576;
Best Local Similarity 99.2%; Pred. No. 3.9e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGGACCACTGTGCAACTTATGCGCGTTTGACAGCGCATAGACCAACAGGG 60
Db 288 ATCAAGCGGACCACTGTGCAACTTATGCGCGTTTGACAGCGCATAGACCAACAGGG 347

Qy 61 GAGTTTATGTTGAATTGTGAGAGATGAGATGTTCTATCTGATCTGCAAGAAGAGG 120
Db 348 GAGTTTATGTTGAATTGTGAGAGATGAGATGTTCTATCTGATCTGCAAGAAGAGG 407

Qy 121 ACCGCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 180
Db 408 ACCGCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 467

Qy 181 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCC-ACCACACT 239
Db 468 GCTAACATTCCTATATTCAGACACAACTGAATACCTTGATCCAGGCTTCCACCACT 527

Qy 240 CAGGCCACCAACCGAT 254
Db 528 CAGGCCACCAACCGAT 542

RESULT 6
E00485
LOCUS E00485 661 bp RNA linear PAT 29-SBP-1997
DEFINITION DNA sequence of a fragment of pSBalpha-3.18.
ACCESSION E00485
VERSION E00485.1 GI:2168768
KEYWORDS JP 1985226888-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
AUTHORS 1 (bases 1 to 661)
Edwards,J.B.D.M., Kerii,B.M., Koorai,H.R. and Henrili,R.
TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE

```

JOURNAL Patent: JP 1985226888-A 3 12-NOV-1985;  
 Cetus Corp  
 COMMENT OS Human (Homo sapiens)  
 PN JP 1985226888-A/3  
 PD 12-NOV-1985  
 PF 20-MAR-1985 JP 1985054705  
 PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI  
 EDWAAADO RUISU SHIRUDON ZA SAADO, KERII BANKUSU MIYURISU, PI  
 KOOREI HAWAADO REBENSON, HENRII RAPOPORTO  
 PC C07D519/00, C07H21/04, C22N15/00, C12O1/68, G01N33/50, G01N33/532,  
 PC G01N33/58,  
 PC (C07D519/00, C07D493/04, C07D495/04), (C07D519/00, C07D493/04, PC  
 C07D493/10);  
 CC strandedness: Double;  
 CC topology: linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: tissue\_type=Blood;  
 CC \*source: cell\_type=B cell;  
 CC \*source: library=cDNA library;  
 EH Key Location/Qualifiers  
 FT misc\_feature 1..661  
 FT /note='a fragment derived from pSBalpha-318  
 FT for insertion'  
 FT Location/Qualifiers  
 FT 1..661  
 FT /organism='unidentified'  
 FT /mol\_type='genomic RNA'  
 FT /db\_xref='taxon:32644'  
 FEATURES source  
 ORIGIN  
 Query Match 95.0%; Score 241.4; DB 6; Length 661;  
 Best Local Similarity 99.2%; Pred. No. 3.9e-63;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 60  
 Db 172 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 231  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 Db 232 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 291  
 QY 121 ACCGTCGGCATCTGGAGGAGTTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGGCGGCTG 180  
 Db 292 ACCGTCGGCATCTGGAGGAGTTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGGCGGCTG 351  
 QY 181 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACT 239  
 Db 352 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACT 411  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 412 CAGGCCACCAACCGAT 426  
 RESULT 8  
 S40633  
 LOCUS S40633 690 bp DNA linear PRI 06-MAY-1993  
 DEFINITION HLA class II: DPAl (DPAl\*0101) [human, Genomic, 690 nt].  
 ACCESSION S40633  
 VERSION S40633.1 GI:1679890  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 690)  
 AUTHORS Marsh, S.G. and Bodmer, J.G.  
 TITLE HLA class II nucleotide sequences, 1991  
 JOURNAL Immunogenetics 33 (5-6), 321-334 (1991)  
 MEDLINE 91267561  
 PUBMED 1904836  
 REMARK GenBank staff at the National Library of Medicine created this  
 entry [NCBI Gibbsq 40633] from the original journal article.  
 COMMENT This sequence comes from Figure 15.  
 On Nov 21, 1996 this sequence version replaced gi:1619630.  
 Region: HLA class II.  
 Location/Qualifiers  
 source 1..690  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 gene 1..690  
 /gene="DPAl"  
 /allele="DPAl\*0101"  
 ORIGIN  
 Query Match 95.0%; Score 241.4; DB 9; Length 690;  
 Best Local Similarity 99.2%; Pred. No. 4e-63;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 60  
 Db 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 60  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 Db 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 JOURNAL Patent: US 4582789-A 5 15-APR-1986;  
 Cetus Corporation; Emeryville, CA  
 Location/Qualifiers  
 JOURNAL Patent: JP 1985226888-A 3 12-NOV-1985;  
 Cetus Corp  
 COMMENT OS Human (Homo sapiens)  
 PN JP 1985226888-A/3  
 PD 12-NOV-1985  
 PF 20-MAR-1985 JP 1985054705  
 PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI  
 EDWAAADO RUISU SHIRUDON ZA SAADO, KERII BANKUSU MIYURISU, PI  
 KOOREI HAWAADO REBENSON, HENRII RAPOPORTO  
 PC C07D519/00, C07H21/04, C22N15/00, C12O1/68, G01N33/50, G01N33/532,  
 PC G01N33/58,  
 PC (C07D519/00, C07D493/04, C07D495/04), (C07D519/00, C07D493/04, PC  
 C07D493/10);  
 CC strandedness: Double;  
 CC topology: linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: tissue\_type=Blood;  
 CC \*source: cell\_type=B cell;  
 CC \*source: library=cDNA library;  
 EH Key Location/Qualifiers  
 FT misc\_feature 1..661  
 FT /note='a fragment derived from pSBalpha-318  
 FT for insertion'  
 FT Location/Qualifiers  
 FT 1..661  
 FT /organism='unidentified'  
 FT /mol\_type='genomic RNA'  
 FT /db\_xref='taxon:32644'  
 FEATURES source  
 ORIGIN  
 Query Match 95.0%; Score 241.4; DB 6; Length 661;  
 Best Local Similarity 99.2%; Pred. No. 3.9e-63;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 60  
 Db 172 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGCGCATAGACCAACAGGG 231  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 Db 232 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 291  
 QY 121 ACCGTCGGCATCTGGAGGAGTTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGGCGGCTG 180  
 Db 292 ACCGTCGGCATCTGGAGGAGTTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGGCGGCTG 351  
 QY 181 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 352 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACT 411  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 412 CAGGCCACCAACCGAT 426  
 RESULT 7  
 I03086  
 LOCUS I03086 661 bp ss-DNA linear PAT 21-MAY-1993  
 DEFINITION Sequence 5 from Patent US 4582789.  
 ACCESSION I03086  
 VERSION I03086.1 GI:268242  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 661)  
 AUTHORS Sheldon, E.L. III, Levenson, C.H., Mullis, K.B. and Rapoport, E.  
 TITLE Process for labeling nucleic acids using psoralen derivatives  
 JOURNAL Patent: US 4582789-A 5 15-APR-1986;  
 Cetus Corporation; Emeryville, CA  
 Location/Qualifiers

QY 121 ACCGCTGGGATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
 Db 121 ACCGCTGGGATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
 QY 181 GCTAACATTGCTATTAATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 181 GCTAACATTGCTATTAATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 240  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 241 CAGGCCACCAACCGAT 255

RESULT 9  
 HUMHDPDA  
 LOCUS Homo sapiens MEC class II DPw3-alpha-1 chain mRNA, linear PRI 07-MAR-1995  
 DEFINITION M27487  
 ACCESSION M27487  
 VERSION M27487.1 GI:703088  
 KEYWORDS cell surface glycoprotein; class II gene; integral membrane protein; lymphocyte antigen; major histocompatibility complex.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 818)  
 AUTHORS Young,J.A., Lindsay,J., Bodmer,J.G. and Trowsdale,J.  
 TITLE Epitope recognition by a DP alpha chain-specific monoclonal antibody (DP11.1) is influenced by the interaction between the DP alpha chain and its polymorphic DP beta chain partner  
 JOURNAL Hum. Immunol. 23 (1), 37-44 (1988)  
 MEDLINE 85053719  
 PUBMED 2461352  
 COMMENT On Mar 9, 1995 this sequence version replaced gi:341719.  
 FEATURES  
 Original Location/Qualifiers  
 source 1..818  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="6p21.3"  
 /cell\_line="LLICRF"  
 /cell\_type="B lymphoblast"  
 gene 1..818  
 /gene="HLA-DPA1"  
 CDS 36..818  
 /gene="HLA-DPA1"  
 /codon\_start=1  
 /product="MHC class II DP3-alpha"  
 /protein\_id="AAA63220.1"  
 /db\_xref="GI:703089"  
 /db\_xref="GDB:G00-120-634"  
 /translation="MRPEDRMFHRAVILRALSLIAFLLSRGAGAIKADHVSTYAAFY QTRPTGRNMFEDDMFYDLDKKEVTWHLFEFGAFAFGAGGLANLAINNNLIT LQSPNLI CHIDKFPFVNLVNCNGELVLPRTDYSHFKHYLTVPFS EGVASLFLPTDYSRKHVLTVPVSAEDFYDCVVEHGLDQLLKHWEAQPIQMP ETVETWLCAGLGLVGLVGIIVGTVLIILKLSRSHDPPRAQGLT"  
 sig\_peptide 36..128  
 /gene="HLA-DPA1"  
 /note="signal peptide fragment"  
 mat\_peptide 129..815  
 /gene="HLA-DPA1"  
 /product="MHC class II DP3-alpha"  
 /note="G00-120-634"

Query Match 95.0%; Score 241.4; DB 9; Length 818;  
 Best Local Similarity 99.2%; Pred. No. 4e-63;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCACCTTATGCGCGGTGTGTACACGCGCATAGACCAAGGG 60

Db 129 ATCAAGCGGACCATGTGTCACCTTATGCGCGGTGTGTACACGCGCATAGACCAAGGG 188  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGATCTGACACGAGAG 120  
 Db 189 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGATCTGACACGAGAG 248  
 QY 121 ACCGCTGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
 Db 249 ACCGCTGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 308  
 QY 181 GCTAACATTGCTATTAATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 309 GCTAACATTGCTATTAATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 368  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 369 CAGGCCACCAACCGAT 383

RESULT 10  
 HSSBAL  
 LOCUS Human mRNA for SB classII histocompatibility antigen alpha-chain. X00457 K01506  
 DEFINITION X00457.1 GI:36405  
 ACCESSION X00457.1  
 VERSION X00457.1  
 KEYWORDS antigen; membrane protein; signal peptide.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1048)  
 AUTHORS Auffray,C., Lillie,J.W., Arnott,D., Grossberger,D., Kappes,D. and Strominger,J.L.  
 TITLE Isotypic and allotypic variation of human class II histocompatibility antigen alpha-chain genes  
 JOURNAL Nature 308 (5957), 327-333 (1984)  
 MEDLINE 84168117  
 PUBMED 6584734  
 COMMENT On Nov 5, 2003 this sequence version replaced gi:188516.  
 FEATURES  
 Location/Qualifiers  
 source 1..1048  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 <1..703  
 /codon\_start=2  
 /product="SB classII histocompatibility antigen alpha-chain"  
 /protein\_id="CAA25143.1"  
 /db\_xref="GI:758100"  
 /db\_xref="GDB:P20036"  
 /db\_xref="SWISS-PROT:P20036"  
 /translation="GAGAIKADHVSTYAAFYQTRPTGRNMFEDDMFYDLDKKE TVWHLFEFGAFAFGAGGLANLAINNNLITLQSPNLI CHIDKFPFVNLVNCNGELVLPRTDYSHFKHYLTVPFS EGVASLFLPTDYSRKHVLTVPVSAEDFYDCVVEHGLDQLLKHWEAQPIQMPETTVLCAGLGLVGLVGIIVGTVLIIL KLSRSHDPPRAQGLT"  
 sig\_peptide <1..13  
 /note="signal peptide fragment"  
 misc\_feature 14..265  
 /note="alpha 1"  
 misc\_feature 266..547  
 /note="alpha 2"  
 misc\_feature 548..586  
 /note="(CP) connecting peptide"  
 misc\_feature 587..655  
 /note="(TW) transmembrane region"  
 misc\_feature 656..700  
 /note="(CY) cytoplasmatic region"  
 polyA\_site 1048  
 /note="polyadenylation site"

ORIGIN

```

Query Match      95.0%; Score 241.4; DB 9; Length 1048;
Best Local Similarity 99.2%; Pred. No. 4.2e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 60
    |||||
Db 14 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 73

QY 61 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 120
    |||||
Db 74 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 133

QY 121 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
    |||||
Db 134 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 193

QY 181 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 239
    |||||
Db 194 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 253

QY 240 CAGGCCACCAACCGAT 254
    |||||
Db 254 CAGGCCACCAACCGAT 268

RESULT 11
103088
LOCUS      1140 bp ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 7 from Patent US 4582789.
ACCESSION 103088
VERSION    103088.1 GI:268244
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1140)
AUTHORS    Sheldon,E.L. III, Levenson,C.H., Mullis,K.B. and Rapoport,H.
TITLE      Process for labeling nucleic acids using psoralen derivatives
JOURNAL    Patent: US 4582789-A 7 15-APR-1986;
            Cetus Corporation; Emeryville, CA
FEATURES   Location/Qualifiers
            source          1..1140
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN
Query Match      95.0%; Score 241.4; DB 6; Length 1140;
Best Local Similarity 99.2%; Pred. No. 4.2e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 60
    |||||
Db 173 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 232

QY 61 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 120
    |||||
Db 233 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 292

QY 121 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
    |||||
Db 293 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 352

QY 181 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 239
    |||||
Db 353 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 412

QY 240 CAGGCCACCAACCGAT 254
    |||||
Db 413 CAGGCCACCAACCGAT 427

RESULT 12
103006
LOCUS      1201 bp ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4617261.
ACCESSION 103006
VERSION    103006.1 GI:268462
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1201)
AUTHORS    Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and
            Watson,R.M.
TITLE      Process for labeling nucleic acids and hybridization probes
JOURNAL    Patent: US 4617261-A 6 14-OCT-1986;
            Cetus Corporation; Emeryville, CA
FEATURES   Location/Qualifiers
            source          1..1201
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN
Query Match      95.0%; Score 241.4; DB 6; Length 1201;
Best Local Similarity 99.2%; Pred. No. 4.2e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 60
    |||||

```

```

LOCUS      103006      1201 bp ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4617261.
ACCESSION 103006
VERSION    103006.1 GI:268462
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1201)
AUTHORS    Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and
            Watson,R.M.
TITLE      Process for labeling nucleic acids and hybridization probes
JOURNAL    Patent: US 4617261-A 6 14-OCT-1986;
            Cetus Corporation; Emeryville, CA
FEATURES   Location/Qualifiers
            source          1..1201
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN
Query Match      95.0%; Score 241.4; DB 6; Length 1201;
Best Local Similarity 99.2%; Pred. No. 4.2e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 60
    |||||
Db 173 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 232

QY 61 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 120
    |||||
Db 233 GAGTTTATGTTGAATTGTGAAGATGAGATGTTCTATGTGGATCTGGACAGAAGGAG 292

QY 121 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
    |||||
Db 293 ACCGTCGGCATCTGGAGAGTTGGCCAAACCTTTCCCTTTGAGGCTCAGGGCGGGCTG 352

QY 181 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 239
    |||||
Db 353 GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACT 412

QY 240 CAGGCCACCAACCGAT 254
    |||||
Db 413 CAGGCCACCAACCGAT 427

RESULT 13
103423
LOCUS      1201 bp ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4822731.
ACCESSION 103423
VERSION    103423.1 GI:270023
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1201)
AUTHORS    Watson,R.M., Sheldon,E.L. III and Sneed,R.M.
TITLE      Process for labeling single-stranded nucleic acids and
            hybridization probes
JOURNAL    Patent: US 4822731-A 6 18-APR-1989;
            Cetus Corporation; Emeryville, CA
FEATURES   Location/Qualifiers
            source          1..1201
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN
Query Match      95.0%; Score 241.4; DB 6; Length 1201;
Best Local Similarity 99.2%; Pred. No. 4.2e-63;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGCGCATAGCCAAACAGGG 60
    |||||

```

Db 173 ATCAAGGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGG 232

QY 61 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 120

Db 233 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 292

QY 121 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180

Db 293 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 352

QY 181 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 239

Db 353 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 412

QY 240 CAGGCCACCAACGAT 254

Db 413 CAGGCCACCAACGAT 427

RESULT 14

AX552229

LOCUS AX552229 1259 bp DNA linear PAT 27-NOV-2002

DEFINITION Sequence 138 from Patent WO0162927.

ACCESSION AX552229

VERSION AX552229.1 GI:25896467

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Banville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K., Amshay, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F., Rosen, B.H., Russo, F.D., Spiro, P.A., Bradley, D.L., Chen, A., Cohen, H.J., Daffo, A., Daniels, S.E., Dufour, G.E., Flores, W., Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S., Roseberry, A.M., Shan, P., Wright, R.J., Yap, P.E., Yu, J.Y., Bratcher, S.R., Chalup, M.S., Dahl, C.R. and Hillman, J.L.

TITLE Polypeptides and corresponding polynucleotides for diagnostics and therapeutics

JOURNAL Patent: WO 0162927-A 138 30-AUG-2001;

FEATURES

source Location/Qualifiers

1..1259

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

/note="Incyte ID No: LI:1169865.1:2000MAY01"

Query Match 95.0%; Score 241.4; DB 6; Length 1259;

Best Local Similarity 99.2%; Pred. No. 4.2e-63;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGG 60

Db 308 ATCAAGGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGG 367

QY 61 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 120

Db 368 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 427

QY 121 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180

Db 428 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 487

QY 181 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 239

Db 488 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 547

QY 240 CAGGCCACCAACGAT 254

Db 548 CAGGCCACCAACGAT 562

RESULT 15

E00484

LOCUS E00484 1201 bp RNA linear PAT 29-SEP-1997

DEFINITION DNA sequence of pSBalpha-318.

ACCESSION E00484

VERSION E00484.1 GI:2168767

KEYWORDS JP 1985226888-A/2.

SOURCE unidentified

ORGANISM unclassified

REFERENCE 1 (bases 1 to 1201)

AUTHORS Edowasdo,R.S.Z.Z.; Kerii,B.M., Koorei,H.R. and Henrii,R.

TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE

JOURNAL Patent: JP 1985226888-A 2 12-NOV-1985;

COMMENT CETUS CORP

OS Human (Homo sapiens)

PN JP 1985226888-A/2

PD 12-NOV-1985

PF 20-MAR-1985 JP 1985054705

PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI

EDOWAADO RUI-SU SHIERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI

KOOREI HAWAADO REBENSON, HENRII RAPOPORUTO

PC C07D519/00,C07H2-/04,C12N15/00,C12Q1/68,G01N33/532,

PC G01N33/58,

PC (C07D519/00,C07D493:04,C07D495:04),(C07D519/00,C07D493:04,PC

C07D493:10);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC \*source: tissue\_type=Blood;

CC \*source: cell\_type=B cell;

CC \*source: library=cdna library;

CC \*source: clone=psBalpha-318;

FH Key Location/Qualifiers

FT CDS 1..2201 /gene="pSBalpha-318"

FT Location/Qualifiers

1..1201

/organism="unidentified"

/mol\_type="genomic RNA"

/db\_xref="taxon:32644"

FEATURES

source

ORIGIN

Query Match 94.4%; Score 239.8; DB 6; Length 1201;

Best Local Similarity 98.8%; Pred. No. 1.3e-62;

Matches 252; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCAAGGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGG 60

Db 173 ATCAAGGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGG 232

QY 61 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 120

Db 233 GAGTTTATGTTGAATTTGATGAAGATGATGTTCTATGTGATCTGACAAAGAGAG 292

QY 121 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180

Db 293 ACCGTTGTCATCTGGAGGAGTTTGCCAAAGCTTTTCCCTTTGAGGCTCAGGCGGGCTG 352

QY 181 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 239

Db 353 GCTAACATTCGTAATTTGAACAACTTTGAATACCTTGATCCAGGTTCCACACACT 412

QY 240 CAGGCCACCAACGAT 254

Db 413 CAGGCCACCAACGAT 427

RESULT 16



HUMHDC02  
LOCUS 2986 bp DNA linear PRI 07-JAN-1995  
DEFINITION Human MHC class II lymphocyte antigen (DPw4-alpha-1) gene, exons 2-4.  
ACCESSION M23904 J02738 M15446  
VERSION M23904.1 GI:188385  
KEYWORDS cell surface glycoprotein; class II gene; integral membrane protein; major histocompatibility complex.  
SEGMENT 2 of 3  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2986)  
AUTHORS Gustafsson,K., Widmark,E., Jonsson,A.K., Serenius,B., Sachs,D.H., Larhammar,D., Rask,L. and Peterson,P.A.  
TITLE Class II genes of the human major histocompatibility complex. Evolution of the DP region as deduced from nucleotide sequences of the four genes  
JOURNAL J. Biol. Chem. 262 (18), 8778-8786 (1987)  
MEDLINE 87250502  
PUBMED 3036929  
COMMENT Original source text: Human T-cell DNA, clone p412-1 and clone p2703-1.  
FEATURES  
source Location/Qualifiers  
1..2986  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
join(M23903.1:563..1095,1..2986)  
/gene="DPw4-alpha-1"  
order(M23903.1:663..1095,1..769)  
/gene="DPw4-alpha-1"  
/number=1  
770..1015  
/gene="DPw4-alpha-1"  
/number=2  
1016..1355  
/gene="DPw4-alpha-1"  
/number=2  
1356..1637  
/gene="DPw4-alpha-1"  
/number=3  
1638..1851  
/gene="DPw4-alpha-1"  
/number=3  
1852..2018  
/gene="DPw4-alpha-1"  
/number=4  
2019..2986  
/gene="DPw4-alpha-1"  
/notes="does not fit consensus"  
/number=4  
ORIGIN About 2.4 kb after segment 1; chromosome 6p21.3.  
Query Match 92.0%; Score 233.8; DB 9; Length 2986;  
Best Local Similarity 98.8%; Pred. No. 1e-60;  
Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 7 GCGACCATCTGTCACACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGAGTTT 66  
Db 769 GCGACCATCTGTCACACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGAGTTT 828  
QY 67 ATGTTTGAATTTGATGATGATGATGATCTTATGTGATCTGGACAAAGGAGACCGTC 126  
Db 829 ATGTTTGAATTTGATGATGATGATGATCTTATGTGATCTGGACAAAGGAGACCGTC 988  
QY 127 TGGCATCTGAGAGTTTGGCCAGCCCTTTTCCTTCAGGCTCAGGCGGCTGGCTAAC 186  
Db 889 TGGCATCTGAGAGTTTGGCCAGCCCTTTTCCTTCAGGCTCAGGCGGCTGGCTAAC 948  
QY 187 ATTGCTATATGACAACTTGATACCTTGATTCAGGTTCC-ACCACTCAGGCC 245  
|||||

Db 949 ATTGCTATATGACAACTTGATACCTTGATTCAGGTTCCACCACTCAGGCC 1008  
QY 246 ACCACCGAT 254  
Db 1009 ACCACCGT 1017  
RESULT 17  
HSHLASBA  
LOCUS HSHLASBA 14646 bp DNA linear PRI 16-FEB-1995  
DEFINITION Human HLA-SB (DP) alpha gene.  
ACCESSION X03100  
VERSION X03100.1 GI:32243  
KEYWORDS antigen; cell surface glycoprotein; class II antigen; glycoprotein; inverted repeat; Kpn repetitive sequence; major histocompatibility complex; repetitive sequence.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 14646)  
AUTHORS Lawrence,S.K., Das,H.K., Pan,J. and Weissman,S.M.  
TITLE The genomic organisation and nucleotide sequence of the HLA-SB (DP) alpha gene  
JOURNAL Nucleic Acids Res. 13 (20), 7515-7528 (1985)  
MEDLINE 86041930  
PUBMED 2997750  
COMMENT Data kindly reviewed (05-MAY-1987) by S.K. Lawrence.  
FEATURES  
source Location/Qualifiers  
1..14646  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
complement(1..247)  
/gene="HLA-SB beta"  
prim\_transcript complement(<1..247)  
intron /gene="HLA-SB beta"  
complement(1..78)  
/gene="HLA-SB beta"  
/number=1  
complement(79..247)  
/gene="HLA-SB beta"  
/number=1  
complement(<79..178)  
/gene="HLA-SB beta"  
/codon\_start=1  
/protein\_id="CAA26886.1"  
/db\_xref="GI:32244"  
/db\_xref="SPTREMBL:O19686"  
/translation="MMVLQVSAAPRTVALTALIMVLTSTVVOGRATP"  
complement(323..374)  
/notes="beta consensus sequence, put. regulatory region"  
514..526  
/repeat\_unit /notes="imp. inverted repeat a"  
2189..2200  
/repeat\_unit /notes="imp. inverted repeat a"  
2461..2512  
/misc\_signal /note="alpha consensus sequence; put. regulatory region"  
2571..11200  
/gene="HLA-SB"  
join(2571..2749,6334..6579,6920..7201,7416..7582,  
10872..11200)  
/gene="HLA-SB alpha"  
join(10872..11200,  
10872..11200)  
/gene="HLA-SB alpha"  
prim\_transcript 2571..11200  
exon /gene="HLA-SB"  
2571..2749  
/number=1  
join(2650..2749,6334..6579,6920..7201,7416..7570)  
/gene="HLA-SB alpha"  
CDS



Query Match 92.0%; Score 233.8; DB 9; Length 64380;  
 Best Local Similarity 98.8%; Pred. No. 1.5e-60;  
 Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 7 GCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGGAGTTT 66  
 Db 2138 GCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGGAGTTT 2079

QY 67 ATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCGATCGACCAAGAGGAGACCGTC 126  
 Db 2078 ATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCGATCGACCAAGAGGAGACCGTC 2019

QY 127 TGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTCAGGCTCAGGCGGCTGGCTTAAC 186  
 Db 2018 TGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTCAGGCTCAGGCGGCTGGCTTAAC 1959

QY 187 ATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACTCAGGCC 245  
 Db 1958 ATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACTCAGGCC 1899

QY 246 ACCACCGAT 254  
 Db 1898 ACCACCGT 1890

RESULT 19  
 AL805913/c  
 LOCUS  
 DEFINITION Human DNA sequence from clone Xxbac-22D21 on chromosome 6, complete sequence.  
 ACCESSION AL805913  
 VERSION AL805913.4 GI:24395073  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Tracey, A.  
 Direct Submission  
 Submitted (23-Oct-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On Oct 25, 2002 this sequence version replaced gi:22204654.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 This sequence was generated from part of bacterial clone configs constructed by the MHC HaploTYPE Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC>

Xxbac-22D21 is from a DNA-arts QBL human bac library VECTOR:  
 pBelOBAC11.  
 FEATURES  
 source Location/Qualifiers  
 1..106728  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXbac-22D21"  
 /clone\_lib="DNA-arts-BAC.1-QBL.1"

## ORIGIN

Query Match 92.0%; Score 233.8; DB 9; Length 106728;  
 Best Local Similarity 98.8%; Pred. No. 1.5e-60;  
 Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 7 GCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGGAGTTT 66  
 Db 95142 GCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGGGAGTTT 95083

QY 67 ATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCGATCGACCAAGAGGAGACCGTC 126  
 Db 95082 ATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCGATCGACCAAGAGGAGACCGTC 95023

QY 127 TGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTCAGGCTCAGGCGGCTGGCTTAAC 186  
 Db 95022 TGGCATCTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTCAGGCTCAGGCGGCTGGCTTAAC 94963

QY 187 ATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACTCAGGCC 245  
 Db 94962 ATTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACTCAGGCC 94903

QY 246 ACCACCGAT 254  
 Db 94902 ACCACCGT 94894

RESULT 20  
 AL845931/c  
 LOCUS  
 DEFINITION Human DNA sequence from clone Xxbac-138A21 on chromosome 6, complete sequence.  
 ACCESSION AL845931  
 VERSION AL845931.7 GI:19572887  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Almeida, J.  
 Direct Submission  
 Submitted (23-Apr-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On Mar 21, 2002 this sequence version replaced gi:19031691.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

REFERENCE  
 1  
 ALMEIDA, J.  
 DIRECT SUBMISSION  
 SUBMITTED (23-APR-2002) WELLCOME TRUST SANGER INSTITUTE, HINXTON, CAMBRIDGESHIRE, CB10 1SA, UK. E-MAIL ENQUIRIES: HUMQUERY@SANGER.AC.UK  
 ON MAR 21, 2002 THIS SEQUENCE VERSION REPLACED GI:19031691.  
 DURING SEQUENCE ASSEMBLY DATA IS COMPARED FROM OVERLAPPING CLONES. WHERE DIFFERENCES ARE FOUND THESE ARE ANNOTATED AS VARIATIONS TOGETHER WITH A NOTE OF THE OVERLAPPING CLONE NAME. NOTE THAT THE VARIATION ANNOTATION MAY NOT BE FOUND IN THE SEQUENCE SUBMISSION CORRESPONDING TO THE OVERLAPPING CLONE, AS WE SUBMIT SEQUENCES WITH ONLY A SMALL OVERLAP AS DESCRIBED ABOVE.  
 THIS SEQUENCE WAS FINISHED AS FOLLOWS UNLESS OTHERWISE NOTED: ALL REGIONS WERE EITHER DOUBLE-STRANDED OR SEQUENCED WITH AN ALTERNATE CHEMISTRY OR COVERED BY HIGH QUALITY DATA (I.E., PHRED QUALITY >= 30); AN ATTEMPT WAS MADE TO RESOLVE ALL SEQUENCING PROBLEMS, SUCH AS COMPRESSIONS AND REPEATS; ALL REGIONS WERE COVERED BY AT LEAST ONE PLASMID SUBCLONE OR MORE THAN ONE M13 SUBCLONE; AND THE ASSEMBLY WAS CONFIRMED BY RESTRICTION DIGEST. THE FOLLOWING ABBREVIATIONS ARE USED TO ASSOCIATE PRIMARY ACCESSION NUMBERS GIVEN IN THE FEATURE TABLE WITH THEIR SOURCE DATABASES: EM.; EMBL; SW.; SWISSPROT; TR.; TREMBL; WP.; WORMPEP; INFORMATION ON THE WORMPEP DATABASE CAN BE FOUND AT

http://www.sanger.ac.uk/Projects/C\_elegans/wormbep/XXbac-138A21 is from a CHORI-501 human bac - PGP cell line library VECTOR: PTARBA2.1

This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/MHC.

## FEATURES

Location/Qualifiers  
1. .124899  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="XXbac-138A21"  
/clone\_lib="CHORI-501"

## ORIGIN

Query Match 92.0%; Score 233.8; DB 9; Length 124899;  
Best Local Similarity 98.8%; Pred. No. 1.6e-60;  
Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
  
QY 7 GCGGACCATGTGTCACTTATGCGGCTTTGTACAGCAGCATAGACCAACAGGGGAGTTT 56  
Db GCGGACCATGTGTCACTTATGCGGCTTTGTACAGCAGCATAGACCAACAGGGGAGTTT 77764  
  
QY 67 AGTTTGAATTTGATGAAGATCAGATGTTCTATGTGGATCTGCAACAAGAGGAGACCGTC 126  
Db AGTTTGAATTTGATGAAGATCAGATGTTCTATGTGGATCTGCAACAAGAGGAGACCGTC 77704  
  
QY 127 TGGCATCTGAGAGTTTGGCAAGCTTTCTTTCAGAGCTCAGGGCGGCTGGCTAAC 186  
Db TGGCATCTGAGAGTTTGGCAAGCTTTCTTTCAGAGCTCAGGGCGGCTGGCTAAC 77644  
  
QY 187 ATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACACTCAGGCC 245  
Db ATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACACTCAGGCC 77584  
  
QY 246 ACCACCGAT 254  
Db ACCACCGAT 77575

## RESULT 21

AC011086/c  
LOCUS  
DEFINITION Homo sapiens chromosome 6 clone RP11-93F3 map 6, WORKING DRAFT  
SEQUENCE, 12 unordered pieces.  
AC011086  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 181228)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lie, C., Locke, K., Macdonald, P., Marquis, N., Morrow, J., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thoman, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.  
2 (bases 1 to 181228)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lie, C., Locke, K., Macdonald, P., Marquis, N., Morrow, J., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thoman, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

## REFERENCE

## AUTHORS

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## Direct Submission

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 181228)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Beda, P., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campobiano, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lie, C., Liu, G., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPreeters, R., Meldrum, J., Meneus, L., Mihova, T., Minnor, T., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thoman, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 10, 2000 this sequence version replaced gi:7717099.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1322

Center clone name: 93\_F3

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator-amersham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 166135 bases at least Q40

Consensus quality: 173448 bases at least Q30

Consensus quality: 176936 bases at least Q20

Insert size: 174000; agarose-fp

Insert size: 180128; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases.

NOTE: This is a working draft sequence. It currently

consists of 12 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence.

as soon as it is available and the accession number will

be preserved.

1 27639: contig of 27639 bp in length

27640 27739: gap of 100 bp

27740 28970: contig of 1231 bp in length

28971 29070: gap of 100 bp

29071 82294: contig of 53224 bp in length

82295 82394: gap of 100 bp

82395 87214: contig of 4820 bp in length

87215 87314: gap of 100 bp

87315 93029: contig of 5715 bp in length

93030 93129: gap of 100 bp

93130 101779: contig of 8650 bp in length

101780 101879: gap of 100 bp

101880 108114: contig of 6235 bp in length

108115 108214: gap of 100 bp

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 08:07:58 ; Search time 236 Seconds  
(without alignments)  
4572.213 Million cell updates/sec

Title: US-09-877-819B-55  
Perfect score: 254  
Sequence: 1 atcaagcgagcatgtgtc.....acactcaggccaccaccgat 254

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 180 summaries

## Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	95.7	255	6	AAL48219 Human HLA
2	241.4	95.0	410	8	ACH49248 Human leu
3	241.4	95.0	436	3	AA43818 Mouse sec
4	241.4	95.0	466	2	AAV86216 EST clone
5	241.4	95.0	476	3	AAC00117 Human sec
6	241.4	95.0	576	3	AAC00116 Human sec
7	241.4	95.0	1202	2	AAQ25060 PSBalpha-
8	241.4	95.0	1259	4	AAQ31123 Human dia
9	241.4	95.0	1348	3	AAF18332 Lung canc
10	238.2	93.8	490	8	ACH49248 Human leu
11	233.8	92.0	14646	6	ABK64796 Human leu
12	230	90.6	362	3	AA43013 Human sec
13	224.4	88.3	396	2	AAV86130 EST clone
14	157.2	61.9	267	2	AAV86047 EST clone
15	146	57.5	267	4	AAK54418 Human hae
16	146	57.5	267	4	AAK54503 Human hae
17	146	57.5	272	4	AAK54317 Human hae
18	146	57.5	294	4	AAK54555 Human hae
19	146	57.5	294	4	AAK54805 Human hae
20	115.6	45.5	641	7	AAQ51219 Human DNA
21	115.6	45.5	747	2	AAV97175 DR alpha-
22	115.6	45.5	750	2	AAV16866 DR-alpha
23	115.6	45.5	750	2	AAV87807 HLA-DR2 a

24	115.6	45.5	819	6	AAI72986 Human HLA
25	115.6	45.5	1183	5	ABA83101 HLA-DR a1
26	115.6	45.5	1446	2	AAT99707 DR2-IgG f
27	115.6	45.5	1446	2	AAK87813 HLA-DR2 a
28	115.6	45.5	1851	2	AAT99708 DR2-IgM z
29	115.6	45.5	1851	2	AAK87814 HLA-DR2 a
30	114	44.9	369	6	ABK27638 Human col
31	114	44.9	399	4	AAL15636 Human bre
32	114	44.9	421	8	ACH29053 Human adu
33	114	44.9	424	8	ACH29217 Human adu
34	114	44.9	444	8	ACH24131 Human adu
35	114	44.9	491	2	AAV89747 EST clone
36	114	44.9	509	3	AAC03760 Human sec
37	114	44.9	542	3	AAC00189 Human sec
38	114	44.9	552	4	ABI99000 Human sin
39	114	44.9	555	4	ABI98999 Human sin
40	114	44.9	558	4	ABI98998 Human sin
41	114	44.9	561	4	ABI99001 Human sin
42	114	44.9	561	4	ABI99003 Human sin
43	114	44.9	561	4	ABI99004 Human sin
44	114	44.9	564	4	ABI99002 Human sin
45	114	44.9	570	4	ABI99006 Human sin
46	114	44.9	570	4	ABI99005 Human sin
47	114	44.9	582	4	ABI99008 Human sin
48	114	44.9	582	4	ABI99007 Human sin
49	114	44.9	600	4	ABI99024 Human sin
50	114	44.9	630	4	ABI99025 Human sin
51	114	44.9	633	4	ABI99012 Human sin
52	114	44.9	636	4	ABI99011 Human sin
53	114	44.9	639	4	ABI99010 Human sin
54	114	44.9	642	4	ABI99016 Human sin
55	114	44.9	642	4	ABI99015 Human sin
56	114	44.9	642	4	ABI99013 Human sin
57	114	44.9	642	4	ABI99026 Human sin
58	114	44.9	645	4	ABI99014 Human sin
59	114	44.9	651	4	ABI99018 Human sin
60	114	44.9	651	4	ABI99017 Human sin
61	114	44.9	663	4	ABI99019 Human sin
62	114	44.9	663	4	ABI99020 Human sin
63	114	44.9	690	2	AAQ92015 DNA encod
64	114	44.9	765	6	AAI19205 Human leu
65	114	44.9	1041	6	AAI19206 DNA encod
66	114	44.9	1099	1	AAK30113 Sequence
67	114	44.9	1099	1	AAK30068 Sequence
68	114	44.9	1099	1	AAK60977 Sequence
69	114	44.9	1199	6	ABK83743 Human cDN
70	114	44.9	1302	3	AAF18036 Lung canc
71	113	44.5	424	3	AAF18365 Lung canc
72	112.4	44.3	243	2	AAT47122 Alpha1 re
73	112.4	44.3	376	4	AAL23750 Human bre
74	112.4	44.3	404	6	ABQ57719 Human col
75	112.4	44.3	588	4	AAL24731 Human bre
76	112.4	44.3	621	2	AAT47123 cDNA enco
77	112.4	44.3	1099	2	AAQ25059 HLA-Dp34
78	110.4	43.5	558	4	ABI98997 Human sin
79	110.4	43.5	639	4	ABI99009 Human sin
80	109.8	43.2	5724	6	ABL69387 Prostate
81	105	41.3	395	4	AAL14891 Human bre
82	103.8	40.9	415	7	ABX36871 Bovine ES
83	101.4	39.9	415	4	AAL24683 Human bre
84	101.4	39.9	175737	6	ABK83571 Human cDN
85	96.4	38.0	600	4	ABI99022 Human sin
86	96.4	38.0	642	4	ABI99023 Human sin
87	95.4	37.6	513	4	AAK54423 Human hae
88	95.4	37.6	514	4	AAK57747 cDNA #423
89	95.4	36.5	433	2	AAV88434 EST clone
90	92.6	36.5	776	2	AAQ03169 Sequence
91	92.6	36.5	776	2	AAQ35054 IAB alpha
92	92.6	36.5	776	2	AAT06285 I-Ab-alpha
93	92.6	36.5	776	2	AAQ56919 Mouse I-A
94	91.8	36.1	180	4	AAL24480 Human bre
95	91.4	36.0	418	7	ABX37815 Bovine ES
96	88	34.6	1243	6	ABN84048 Single ch

97	87.6	34.5	261	2	AAT60684	Alpha re	Aat60684	Alpha re	170	69	27.2	772	3	AAFL18248	Lung canc
98	87.6	34.5	588	2	AAT60698	Alphalalp	Aat60698	Alphalalp	171	69	27.2	1261	6	ABJ67545	Thyroid c
99	87.6	34.5	654	2	AAT60698	cdna enco	Aat60698	cdna enco	172	68.4	26.9	1821	6	ABQ54583	Human inn
100	87.6	34.5	1344	2	AAT60705	cdna enco	Aat60705	cdna enco	173	67.4	26.5	580	4	AAS28875	Human inn
101	87.6	34.5	1382	2	AAT60705	cdna enco	Aat60705	cdna enco	174	67.4	26.5	580	4	ABA06612	Human cdn
102	87.6	34.5	1382	2	AAT60705	cdna enco	Aat60705	cdna enco	175	67.4	26.5	580	4	ABV83949	Human pol
103	87.6	34.5	1382	2	AAT60705	cdna enco	Aat60705	cdna enco	176	67.4	26.5	580	4	ADB31600	Human cdn
104	87.6	34.5	1385	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC	177	65.6	25.8	275	2	AAV89807	EST clone
105	87.6	34.5	1385	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC	178	64.6	25.4	585	4	AAK54616	Human hae
106	87.6	34.5	1385	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC	179	64.6	25.4	585	4	AAK54458	Human hae
107	87.6	34.5	1484	5	AAFP5098	cdna enco	Aaf55098	DNA enco	180	64.6	25.4	585	4	AAK54468	Human hae
108	87.6	34.5	1508	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC							
109	87.6	34.5	1508	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC							
110	87.6	34.5	1508	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC							
111	87.6	34.5	1508	2	AAT60744	Mouse MHC	Aat60744	Mouse MHC							
112	87.6	34.5	4713	2	AAV12067	Marine IA	Aav12067	Marine IA							
113	87.6	34.3	106	1	AAV12067	Marine IA	Aav12067	Marine IA							
114	87.2	34.3	106	1	AAV12067	Marine IA	Aav12067	Marine IA							
115	86.8	34.3	106	1	AAV12067	Marine IA	Aav12067	Marine IA							
116	86	33.9	2481	7	ABT41766	Toxicity	Abt41766	Toxicity							
117	85.4	33.6	405	4	ABA09442	Human HLA	Aba09442	Human HLA							
118	84.2	33.1	261	2	AAT60685	Alphalalp	Aat60685	Alphalalp							
119	84.2	33.1	588	2	AAT60701	Alphalalp	Aat60701	Alphalalp							
120	84.2	33.1	609	4	ABJ99034	Marine pC	Abj99034	Marine pC							
121	84.2	33.1	614	4	ABJ99035	Marine pC	Abj99035	Marine pC							
122	84.2	33.1	633	2	AAT60688	cdna enco	Aat60688	cdna enco							
123	84.2	33.1	642	4	ABJ99036	Marine pC	Abj99036	Marine pC							
124	84.2	33.1	647	4	ABJ99037	Marine pC	Abj99037	Marine pC							
125	84.2	33.1	702	2	AAT60704	cdna enco	Aat60704	cdna enco							
126	84.2	33.1	702	2	ABJ99043	Marine pC	Abj99043	Marine pC							
127	84.2	33.1	773	4	ABJ99042	Marine pC	Abj99042	Marine pC							
128	84.2	33.1	1013	4	ABJ99044	Marine pC	Abj99044	Marine pC							
129	84.2	33.1	1223	2	AAT60700	cdna enco	Aat60700	cdna enco							
130	84.2	33.1	1662	4	ABJ99039	Marine pC	Abj99039	Marine pC							
131	84.2	33.1	1676	4	ABJ99041	Marine pC	Abj99041	Marine pC							
132	84.2	33.1	1680	4	ABJ99021	Marine pC	Abj99021	Marine pC							
133	84.2	33.1	1866	4	ABJ99031	MBP 1-14	Abj99031	MBP 1-14							
134	84.2	33.1	1866	4	ABJ99031	MBP 1-14	Abj99031	MBP 1-14							
135	84.2	33.1	1898	4	ABJ99038	Marine pC	Abj99038	Marine pC							
136	84.2	33.1	1701	4	ABJ99028	IAS MBP 1	Abj99028	IAS MBP 1							
137	84.2	33.1	2053	4	ABJ99029	IAS MBP 9	Abj99029	IAS MBP 9							
138	84.2	33.1	2059	4	ABJ99032	MBP 1-14	Abj99032	MBP 1-14							
139	84.2	33.1	2343	4	ABJ99033	MBP 90-10	Abj99033	MBP 90-10							
140	83.4	32.8	2346	4	ABJ99027	IAS MBP 1	Abj99027	IAS MBP 1							
141	83.4	32.8	956	8	AD548319	Human pol	Aai58351	Human pol							
142	83.4	32.8	977	4	AAH98341	Human EST	Aah98341	Human EST							
143	83.4	32.8	979	4	AAI60137	Human pol	Aai60137	Human pol							
144	83.4	32.8	995	4	ABA06355	Human cdn	Aba06355	Human cdn							
145	83.4	32.8	995	6	ABV83692	Human pol	Abv83692	Human pol							
146	83.4	32.8	2597	3	AAQ93313	Human sec	Aac93313	Human sec							
147	82.8	32.6	566	2	AAQ93313	Nucleic a	Aax32523	Nucleic a							
148	82.8	32.6	566	7	AAQ93313	Nucleic a	Aax32523	Nucleic a							
149	82.2	32.4	323	7	ABX44671	Bovine BS	Abx44671	Bovine BS							
150	81.8	32.2	242	2	AAQ25434	HLA Class	Aaq25434	HLA Class							
151	81.8	32.2	242	2	AAQ25434	HLA Class	Aaq25434	HLA Class							
152	81.8	32.2	242	2	AAQ25434	HLA Class	Aaq25434	HLA Class							
153	81.2	32.0	564	4	AAI15887	Human bre	Aai15887	Human bre							
154	80.4	31.7	434	3	AAA45351	Human sec	Aaa45351	Human sec							
155	80.4	31.7	441	8	ACH27689	Human adu	Ach27689	Human adu							
156	80.4	31.7	512	7	ACD95045	Human col	Acd95045	Human col							
157	80.4	31.7	768	6	ABQ74306	Human leu	Abq74306	Human leu							
158	80.4	31.7	768	6	ABQ74306	Human leu	Abq74306	Human leu							
159	80.4	31.7	1060	6	AAQ25064	Human leu	Aaq25064	Human leu							
160	80.4	31.7	1060	6	AAQ25064	Human leu	Aaq25064	Human leu							
161	79.4	31.3	1020	4	AAQ25064	Human leu	Aaq25064	Human leu							
162	79.4	31.3	1020	4	AAQ25064	Human leu	Aaq25064	Human leu							
163	77.2	30.4	467	9	ACH21926	Human adu	Ach21926	Human adu							
164	76.4	30.1	459	4	AAI09097	Human bre	Aai09097	Human bre							
165	74.8	29.4	193	4	AAI09097	Human bre	Aai09097	Human bre							
166	73	28.7	1707	4	ABJ99030	IAS MBP 9	Abj99030	IAS MBP 9							
167	72.2	28.4	373	2	AAV86139	EST clone	Aav86139	EST clone							
168	72.2	28.4	3001	6	ABK34025	Human DNA	Abk34025	Human DNA							
169	71.4	28.1	440	3	AAQ04158	Human sec	Aac04158	Human sec							

## ALIGNMENTS

## RESULT 1

AAFL48219  
ID AAL48219 standard; DNA; 255 BP.

XX AAL48219;  
AC AAL48219;

DT 01-OCT-2002 (first entry)

XX Human HLA DPA1 exon 2 sequence.

DE Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
KW flow cytometry; human; DPA1; DPB1; Gene; ds.  
KW Homo sapiens.

OS WO200194639-A1.

XX 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US018590.

XX 08-JUN-2000; 2000US-0210759p.

XX (REGC) UNIV CALIFORNIA.

XX White PS, Torney DC;

XX WPI; 2002-566450/60.

XX Identifying sequences useful as address/capture tags for flow cytometry based minisequencing, by generating tag sequences and rejecting sequences based on certain parameters e.g. sequences which form stable hairpins.

XX Disclosure; Fig 4; 35pp; English.

XX The present invention relates to a method of identifying sequences useful as address/capture tags, involving rejecting sequences having common sub-sequences with a sub-sequence length greater than specified number of bases, and sequences which can form stable hairpins and stable dimers from a sample of oligonucleotides, and selecting those sequences in the sample that would hybridise to their respective complements with a high degree of specificity. The method is useful for identifying a set of sequences useful as address/capture tags which can be used for multiplexed single nucleotide polymorphism (SNP) scoring in a flow cytometry assay. The present sequence is a fragment of the human DPA1 gene described in the exemplification of the invention

XX Sequence 255 BP; 67 A; 57 C; 66 G; 65 T; 0 U; 0 Other;

Query Match 95.7%; Score 243; DB 6; Length 255;

Best Local Similarity 99.6%; Pred. No. 3.4e-67;

Matches 254; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 ATCAAGCGGACCATGTGTCACTTATGCGCGTTTGTACAGCGCATACACACAGGG 60

Db 1 ATCAAGCGGACCATGTGTCACTTATGCGCGTTTGTACAGCGCATACACACAGGG 60

```
QY 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGCTTCTATGCTGATCTGGCAAGAGGAG 120
Db 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGCTTCTATGCTGATCTGGCAAGAGGAG 120
QY 121 ACGTCTGGCACTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGGCGGCTG 180
Db 121 ACGTCTGGCACTGAGGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGGCGGCTG 180
QY 181 GCTTACATGCTATATGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACT 239
Db 181 GCTTACATGCTATATGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACT 240
QY 240 CAGGCCACCAACCGAT 254
Db 241 CAGGCCACCAACCGAT 255

RESULT 2
ACH49248
ID ACH49248 standard; cDNA; 410 BP.
XX AC ACH49248;
XX AC ACH49248;
XX DT 13-OCT-2003 (first entry)
XX Human leukocyte cDNA #842.
XX DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX PS WPI; 2003-615964/58.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH42789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
```

```
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX SQ Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;
Query Match 95.0%; Score 241.4; DB 8; Length 410;
Best Local Similarity 99.2%; Pred. No. 1.3e-66;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATCAAGCGGACCATGTGCAACTTATGCGCGGTTTGTACAGACGCATAGACCAAGG 60
Db 145 ATCAAGCGGACCATGTGCAACTTATGCGCGGTTTGTACAGACGCATAGACCAAGG 204
QY 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGCTTCTATGCTGATCTGGCAAGAGGAG 120
Db 205 GAGTTTATGTTGAAATTTGATGAGATGAGATGCTTCTATGCTGATCTGGCAAGAGGAG 264
QY 121 ACCGCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGGCGGCTG 180
Db 265 ACCGCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGGCGGCTG 324
QY 181 GCTTACATGCTATATGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACT 239
Db 325 GCTTACATGCTATATGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACT 384
QY 240 CAGGCCACCAACCGAT 254
Db 385 CAGGCCACCAACCGAT 399

RESULT 3
AAA43818
ID AAA43818 standard; cDNA; 436 BP.
XX AC AAA43818;
XX AC AAA43818;
XX DT 21-AUG-2000 (first entry)
XX DE Mouse secreted expressed sequence tag SEQ ID NO:393.
XX KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
XX expressed sequence tag; EST; probe; chemotactic; proliferative;
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
XX antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
XX autoimmune disorder; multiple sclerosis; allergic condition;
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
XX central nervous system disorder; Alzheimer's disease; stroke;
XX Parkinson's disease; Huntington's disease; coagulation disorder;
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
XX infection; depression; psoriasis; ss.
XX OS Mus musculus.
XX EN WO200021991-A1.
XX PD 20-APR-2000.
XX PF 15-OCT-1999; 99WO-US024206.
XX PR 15-OCT-1998; 98US-0104436P.
XX PA (GENY ) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Bowman MR;
XX WPI; 2000-317938/27.
XX Isolated polynucleotides, and encoded proteins, comprising secreted
XX expressed sequence tags (SESTs), useful for treating various disorders
XX PT
```

PT such as autoimmune, infectious, and central nervous system disorders.  
 PS Claim 1; Page 306; 803pp; English.  
 XX  
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue  
 CC sources. The sESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; anti-inflammatory; cytostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antischismatic; vulnary; antitumor;  
 CC osteoprotective; neuroprotective; neurotropic; antiparkinsonian; antipsoriatic;  
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention  
 XX

SQ Sequence 436 BP; 108 A; 113 C; 106 G; 109 T; 0 U; 0 Other;

Query Match 95.0%; Score 241.4; DB 3; Length 436;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-66;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCCGCTTTGTACACGATAGACCAACAGG 60  
 DB 167 ATCAAGCGGACCATGTGTCAACTTATGCCGCTTTGTACACGATAGACCAACAGG 226  
 QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGCAAGAAGGAG 120  
 DB 227 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGCAAGAAGGAG 286  
 QY 121 ACCGCTCGCATCTGAGAGTTTGGCAAGCTTTCTCTTTGAGGCTCAGGCGGGCTG 180  
 DB 287 ACCGCTCGCATCTGAGAGTTTGGCAAGCTTTCTCTTTGAGGCTCAGGCGGGCTG 346  
 QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCACGCTTCC-ACCACACT 239  
 DB 347 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCACGCTTCC-ACCACACT 406  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 407 CAGGCCACCAACCGAT 421

RESULT 4

AAV86216

ID AAV86216 standard; cDNA; 466 BP.

XX AC AAV86216;

XX AC AAV86216;

DT 27-APR-1999 (first entry)

XX EST clone O67.

XX Expressed sequence tag; secreted protein; haematopoiesis regulator;

XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;

XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;

XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX Homo sapiens.

OS

XX

XX WO9845435-A2.

PN

XX

PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US006954.  
 XX  
 PR 10-APR-1997; 97JS-00835913.  
 XX  
 PA (GEMY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCooy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 PI Spaulding V, Agostino MJ;  
 PI WPI; 1999-070076/06.  
 XX

New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 PT

XX Claim 1; Page 160-161; 633pp; English.

XX This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX

SQ Sequence 466 BP; 110 A; 122 C; 122 G; 112 T; 0 U; 0 Other;

Query Match 95.0%; Score 241.4; DB 2; Length 466;

Best Local Similarity 99.2%; Pred. No. 1.4e-66;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCCGCTTTGTACACGATAGACCAACAGG 60  
 DB 113 ATCAAGCGGACCATGTGTCAACTTATGCCGCTTTGTACACGATAGACCAACAGG 172  
 QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGCAAGAAGGAG 120  
 DB 173 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGCAAGAAGGAG 232  
 QY 121 ACCGCTCGCATCTGAGAGTTTGGCAAGCTTTCTCTTTGAGGCTCAGGCGGGCTG 180  
 DB 233 ACCGCTCGCATCTGAGAGTTTGGCAAGCTTTCTCTTTGAGGCTCAGGCGGGCTG 292  
 QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCACGCTTCC-ACCACACT 239  
 DB 293 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCACGCTTCC-ACCACACT 352  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 353 CAGGCCACCAACCGAT 367

RESULT 5

AAAC00117

ID AAC00117 standard; cDNA; 476 BP.

XX AC AAC00117;

XX AC AAC00117;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 115.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

XX



```

OS Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00111.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 115; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors
XX
XX Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;
XX
XX Query Match 95.0%; Score 241.4; DB 3; Length 476;
XX Best Local Similarity 99.2%; Pred. No. 1.4e-66;
XX Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
XX
XX 1 ATCAAGGCGGACCATGTGTCAACTTATCCCGTTTGTACAGACGATAGACCAACAGGG 60
XX 188 ATCAAGGCGGACCATGTGTCAACTTATCCCGTTTGTACAGACGATAGACCAACAGGG 247
XX
XX 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTTGGATCTGGACAAGAGGAG 120
XX 248 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTTGGATCTGGACAAGAGGAG 307
XX
XX 121 ACCGTCCTGGATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
XX 308 ACCGTCCTGGATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 367
XX
XX 181 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGTACCGCTTCC-ACCACACT 239
XX 368 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTTCCACACACT 427
XX
XX 240 CAGGCCACCAACCGAT 254
XX 428 CAGGCCACCAACCGAT 442
XX
XX RESULT 6
XX AAC00116
XX ID AAC00116 standard; cDNA; 576 BP.
XX
XX AC AAC00116;
XX
XX 06-OCT-2000 (first entry)
XX
XX
XX

```

```

DE Human secreted protein 5' EST, SEQ ID NO: 114.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00110.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 114; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors
XX
XX Sequence 576 BP; 131 A; 158 C; 141 G; 145 T; 0 U; 1 Other;
XX
XX Query Match 95.0%; Score 241.4; DB 3; Length 576;
XX Best Local Similarity 99.2%; Pred. No. 1.6e-66;
XX Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
XX
XX 1 ATCAAGGCGGACCATGTGTCAACTTATCCCGTTTGTACAGACGATAGACCAACAGGG 60
XX 288 ATCAAGGCGGACCATGTGTCAACTTATCCCGTTTGTACAGACGATAGACCAACAGGG 347
XX
XX 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTTGGATCTGGACAAGAGGAG 120
XX 348 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTTGGATCTGGACAAGAGGAG 407
XX
XX 121 ACCGTCCTGGATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
XX 408 ACCGTCCTGGATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 467
XX
XX 181 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTTCC-ACCACACT 239
XX 468 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTTCCACACACT 527
XX
XX 240 CAGGCCACCAACCGAT 254
XX 528 CAGGCCACCAACCGAT 542
XX
XX RESULT 7
XX AAQ25060
XX ID AAQ25060 standard; DNA; 1202 BP.

```

XX AAQ25060;  
 XX AC  
 XX DT 15-NOV-1992 (first entry)  
 XX DE  
 XX DE pSbalpha-318 clone.  
 XX KW RFLP; probe p29G8; beta-lymphoblastoid cell line; LG2; T5-1; 6.3.6;  
 XX KW transplant; transfection; paternity; ss.  
 XX OS Synthetic.  
 XX PN US5110920-A.  
 XX PD 05-MAY-1992.  
 XX PF 05-DEC-1984; 84US-00678255.  
 XX PR 22-JAN-1982; 82US-00341902.  
 XX PR 07-JAN-1983; 83US-00456373.  
 XX PR 30-AUG-1988; 88US-00238619.  
 XX PA (CETU ) CETUS CORP.  
 XX PI Erlich HA;  
 XX DR WPI; 1992-175244/21.  
 XX PT New DNA probes specific to single class II HLA locus - useful in HLA  
 PT typing e.g. to evaluate paternity and transplant or transfusion  
 PT compatibility and to diagnose disease susceptibility.  
 XX PS Disclosure; Page 11; 21pp; English.  
 XX CC The sequence given is a pSbalpha-318 clone which was derived from a beta-  
 CC lymphoblastoid cell line IG2 cDNA library using a probe designated p29G8.  
 CC This probe bound to sequences distinct from those which lead to the  
 CC elucidation of HLA-Dp34 (see also AAQ25059). A genomic blot pattern with  
 CC DNA from the cell lines T5-1 and its HLA hemizygous derivative 6.3.6  
 CC indicates that the p29G8 locus maps within the HLA region. p29G8 has been  
 CC found to be a HLA-Sbalpha clone and could be used to isolate the given  
 CC sequence. p29G8 could be useful in HLA typing based on RFLPs. It can be  
 CC utilized in paternity disputes or for determining transplant or transfusion  
 CC compatibility. It can also be used to make disease correlations to  
 CC diagnose diseases or predict susceptibility to diseases  
 XX SQ Sequence 1202 BP; 299 A; 336 C; 277 G; 300 T; 0 U; 0 Other; .  
 Query Match 95.0%; Score 241.4; DB 2; Length 1202;  
 Best Local Similarity 99.2%; Pred. No. 2.1e-66;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTCTCAACTATATCCCGTTTGTACAGCCATAGACCAACAGGG 60  
 DB 173 ATCAAGCGGACCATGTCTCAACTATATCCCGTTTGTACAGCCATAGACCAACAGGG 232  
 QY 61 GAGTTTATGTTGATGAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 DB 233 GAGTTTATGTTGATGAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 292  
 QY 121 ACCGTCTGGCATCTGAGAGAGTTGGCAAGCCTTTTCTTTGAGGCTCAGGGGGGGGTG 180  
 DB 293 ACCGTCTGGCATCTGAGAGAGTTGGCAAGCCTTTTCTTTGAGGCTCAGGGGGGGGTG 352  
 QY 181 GCTACATGCTATATTGAACAACACTGTATACCTTGATCCAGCTTCC-ACCACACT 239  
 DB 353 GCTACATGCTATATTGAACAACACTGTATACCTTGATCCAGCTTCCACCAACACT 412  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 413 CAGGCCACCAACGAT 427

RESULT 8  
 RAS31123  
 ID AAS31123 standard; cDNA; 1259 BP.  
 XX AC  
 XX AC AAS31123;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE Human diagnostic and therapeutic polynucleotide (DITHP) #138.  
 XX KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
 XX KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
 XX KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
 XX KW respiratory disorder; ss.  
 XX OS Homo sapiens.  
 XX PN WO200162927-A2.  
 XX PD 30-AUG-2001.  
 XX PF 21-FEB-2001; 2001WO-US006059.  
 XX PR 24-FEB-2000; 2000US-0184693P.  
 XX PR 24-FEB-2000; 2000US-0184697P.  
 XX PR 24-FEB-2000; 2000US-0184698P.  
 XX PR 24-FEB-2000; 2000US-0184768P.  
 XX PR 24-FEB-2000; 2000US-0184769P.  
 XX PR 24-FEB-2000; 2000US-0184770P.  
 XX PR 24-FEB-2000; 2000US-0184771P.  
 XX PR 24-FEB-2000; 2000US-0184772P.  
 XX PR 24-FEB-2000; 2000US-0184773P.  
 XX PR 24-FEB-2000; 2000US-0184774P.  
 XX PR 24-FEB-2000; 2000US-0184776P.  
 XX PR 24-FEB-2000; 2000US-0184777P.  
 XX PR 24-FEB-2000; 2000US-0184797P.  
 XX PR 24-FEB-2000; 2000US-0184813P.  
 XX PR 24-FEB-2000; 2000US-0184837P.  
 XX PR 24-FEB-2000; 2000US-0184841P.  
 XX PR 24-FEB-2000; 2000US-0185213P.  
 XX PR 24-FEB-2000; 2000US-0185216P.  
 XX PR 12-MAY-2000; 2000US-0203785P.  
 XX PR 15-MAY-2000; 2000US-0204226P.  
 XX PR 16-MAY-2000; 2000US-0204525P.  
 XX PR 16-MAY-2000; 2000US-0204821P.  
 XX PR 16-MAY-2000; 2000US-0204908P.  
 XX PR 16-MAY-2000; 2000US-0205232P.  
 XX PR 17-MAY-2000; 2000US-0204815P.  
 XX PR 17-MAY-2000; 2000US-0204863P.  
 XX PR 17-MAY-2000; 2000US-0205221P.  
 XX PR 17-MAY-2000; 2000US-0205285P.  
 XX PR 17-MAY-2000; 2000US-0205286P.  
 XX PR 17-MAY-2000; 2000US-0205287P.  
 XX PR 17-MAY-2000; 2000US-0205323P.  
 XX PR 17-MAY-2000; 2000US-0205324P.  
 XX PA (INCY-) INCYTE GENOMICS INC.  
 XX PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
 XX PI Chen A, D'sa SA, Amsberg S, Dahl CR, Dam TC, Daniels SE, Dufour GE;  
 XX PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TF;  
 XX PI Roseberry AM, Rosen BH, Russo PD, Stockdreher TK, Daffo A;  
 XX PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
 XX PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
 XX DR WPI; 2001-502867/55.  
 XX DR P-ESDE; AAU19552.  
 XX PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 XX PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 XX PS Claim 1; Page 361; 522pp; English.  
 XX XX

CC The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DITHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide  
 CC expression by rectifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DITHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DITHPs and in assays to identify modulators of DITHP  
 CC expression and activity. The anti-DITHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DITHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay  
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic  
 CC (DITHP) polynucleotides of the invention

XX Sequence 1259 BP; 293 A; 350 C; 300 G; 316 T; 0 U; 0 Other;

Query Match 95.0%; Score 241.4; DB 4; Length 1259;  
 Best Local Similarity 99.2%; Pred. No. 2.2e-66;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTCTCACTATGCGCGCTTGTACAGACGCATAGACCAACAGG 60  
 DB 308 ATCAAGCGGACCATGTCTCACTATGCGCGCTTGTACAGACGCATAGACCAACAGG 367  
 QY 61 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGCGATCTGCAAGAGAG 120  
 DB 368 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGCGATCTGCAAGAGAG 427  
 QY 121 ACCGTCGCGCATCTGGAGGAGTTTGGCCAAACGCTTTCTTTGAGGCTCAGGCGGGCTG 180  
 DB 428 ACCGTCGCGCATCTGGAGGAGTTTGGCCAAACGCTTTCTTTGAGGCTCAGGCGGGCTG 487  
 QY 181 GCTAACATTCGTATATGAACAACAACCTGTAATACCTTGATCCGCGCTTCC-ACCACT 239  
 DB 488 GCTAACATTCGTATATGAACAACAACCTGTAATACCTTGATCCGCGCTTCCACCACT 547  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 548 CAGGCCACCAACCGAT 562

RESULT 9  
 AAF18332  
 ID AAF18332 standard; DNA; 1348 BP.  
 XX AAF18332;  
 AC AC  
 AC AC  
 XX 14-MAR-2001 (first entry)  
 DT 14-MAR-2001 (first entry)  
 XX Lung cancer associated polynucleotide sequence SEQ ID 351.  
 DE DE  
 XX Human; lung cancer associated protein; neuroprotective; cytosolic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerable;  
 KW gastrointestinal; nephrotropic; antiinfective; synecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.  
 XX Homo sapiens.  
 OS OS  
 XX WC200055180-A2.  
 PN PN  
 XX

PD 21-SEP-2000.  
 XX 08-MAR-2000; 2000WO-US005918.  
 PF 12-MAR-1999; 99US-0124270P.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 XX (ROSE/) ROSEN C A.  
 PA Ruben SM;  
 XX WPI; 200Q-587514/55.  
 XX P-PSDB; AAB58456.  
 DR Lung cancer associated gene sequences, referred to as lung cancer  
 XX PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer.

XX Claim 1; Page 808-809; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytosolic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerable; gastrointestinal  
 CC general; nephrotropic; antiinfective; synecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the protein  
 CC or polynucleotide sequences. The lung cancer associated polynucleotide  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The proteins may be used to treat disorders such as  
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,  
 CC cardiovascular, renal, and proliferative disorders. The proteins may also  
 CC be used in the treatment of wounds and infectious diseases.  
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are  
 CC used in the course of the invention for the identification and  
 CC characterisation of the polynucleotide and protein sequences

XX Sequence 1348 BP; 335 A; 373 C; 307 G; 324 T; 0 U; 9 Other;

Query Match 95.0%; Score 241.4; DB 3; Length 1348;  
 Best Local Similarity 99.2%; Pred. No. 2.2e-66;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTCTCACTATGCGCGCTTGTACAGACGCATAGACCAACAGG 60  
 DB 246 ATCAAGCGGACCATGTCTCACTATGCGCGCTTGTACAGACGCATAGACCAACAGG 305  
 QY 61 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGCGATCTGCAAGAGAG 120  
 DB 306 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGCGATCTGCAAGAGAG 365  
 QY 121 ACCGTCGCGCATCTGGAGGAGTTTGGCCAAACGCTTTCTTTGAGGCTCAGGCGGGCTG 180  
 DB 366 ACCGTCGCGCATCTGGAGGAGTTTGGCCAAACGCTTTCTTTGAGGCTCAGGCGGGCTG 425  
 QY 181 GCTAACATTCGTATATGAACAACAACCTGTAATACCTTGATCCAGCGTTC-ACCACT 239  
 DB 426 GCTAACATTCGTATATGAACAACAACCTGTAATACCTTGATCCAGCGTTCACCACT 485  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 486 CAGGCCACCAACCGAT 500

RESULT 10  
 ACH49928  
 ID ACH49928 standard; cDNA; 490 BP.  
 XX ACH49928;  
 AC AC  
 XX 13-OCT-2003 (first entry)  
 DT 13-OCT-2003 (first entry)  
 XX

DE Human leukocyte cDNA #1522.  
XX Human; ss; sequencing by hybridisation; SSH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
OS US2003073623-A1.  
XX 17-APR-2003.  
PD 30-JUL-2001; 2001US-00918995.  
XX 30-JUL-2001; 2001US-00918995.  
XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX Claim 1; SEQ ID NO 37140; 44pp; English.  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SSH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX Sequence 490 BP; 112 A; 130 C; 125 G; 121 T; 0 U; 2 Other;  
Query Match 93.8%; Score 238.2; DB 8; Length 490;  
Best Local Similarity 98.4%; Pred. No. 1.5e-65;  
Matches 251; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Q/ 1 ATCAAGCGGACCATGTGTCACTTATGCCCGTTTGTACAGCGCATAGACCAAGGG 60  
D/ 218 ATCAAGCGGACCATGTGTCACTTATGCCCGTTTGTACAGCGCATAGACCAAGGG 277  
Q/ 61 GAGTTTATGTTGATTTGATGAGATGAGATGTTCTATGTGGATCTGGACAAGAGAG 120  
D/ 278 GAGTTTATGTTGATTTGATGAGATGAGATGTTCTATGTGGATCTGGACAAGAGAG 337  
Q/ 121 ACCGTCTGGCATCTGGAGGAGTTTGGCAAGCCTTTTCCCTTTGAGGCTCAGGGGGGCTG 180  
D/ 338 ACCGTCTGGCATCTGGAGGAGTTTGGCAAGCCTTTTCCCTTTGAGGCTCAGGGGGGCTG 397  
Q/ 181 GCTAACATTGCTATTGAAACAACTTGAATACCTTGATCCAGCTTC-ACCACACT 239  
D/ 398 GCTAACATTGCTATTGAAACAACTTGAATACCTTGATCCAGCTTC-ACCACACT 457  
Q/ 240 CAGGCCCAACCCGAT 254

Db 458 CAGGCCCAACCCGAT 472  
RESULT 11  
ABK64796  
ID ABK64796 standard; DNA; 14646 BP.  
XX AC ABK64796;  
XX DT 18-JUN-2002 (first entry)  
DE Human benign prostatic hyperplasia gene #691.  
XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
XX Homo sapiens.  
XX WO200212440-A2.  
XX PD 14-FEB-2002.  
XX PF 07-AUG-2001; 2001WO-US024708.  
XX PR 07-AUG-2000; 2000US-0223323P.  
XX PR 05-JUN-2001; 2001US-00873319.  
XX PA (GENE-) GENE LOGIC INC.  
XX PA (NLSB) JAPAN TOBACCO INC.  
XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
XX WPI; 2002-257476/30.  
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
PT detecting expression levels of one or more genes in prostate cells from  
PT patient that are differentially regulated compared to normal prostate  
XX cells.  
XX Disclosure; Page 373-377; 44pp; English.  
XX The invention relates to a method of diagnosing (I) the onset or  
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
CC or identifying an agent that modulates the onset or progression of BPH.  
CC The method is based on changes in gene expression in BPH tissue isolated  
CC from patients exhibiting different clinical states of prostate  
CC hyperplasia as compared to normal prostate tissue. (I) comprises  
CC detecting the expression levels of one or more genes in prostate cells  
CC from the subject that are differentially regulated compared to normal  
CC of BPH cells or BPH-like cell population, exposing the cells to the  
CC agent, preparing a second gene expression profile of the agent exposed  
CC cells, and comparing the first and second gene expression profiles. (I)  
CC is useful for diagnosing the onset or progression of BPH. (II) is useful  
CC for identifying an agent that modulates the onset or progression of BPH.  
CC The methods are useful to present information identifying the expression  
CC level in a tissue or cells, by comparing the expression level of genes  
CC given in the specification in the tissue or cells to the level of  
CC expression of gene in the database, and displaying the expression levels  
CC of at least one gene in the tissue or cell sample compared to the  
CC expression level in BPH. Agents using (II) are useful for treating BPH or  
CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic  
CC hyperplasia gene sequences of the invention  
XX SQ Sequence 14646 BP; 4552 A; 2957 C; 2741 G; 4395 T; 0 U; 1 Other;  
Query Match 92.0%; Score 233.8; DB 6; Length 14646;  
Best Local Similarity 98.8%; Pred. No. 1.1e-63;  
Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
Q/ 7 GCGGACCATGTGTCACTTATGCCCGTTTGTACAGCGCATAGACCAAGGGAGTTT 66  
D/ 6333 GCGGACCATGTGTCACTTATGCCCGTTTGTACAGCGCATAGACCAAGGGAGTTT 6392



PT pituitary, retina and colon cDNA libraries.  
 XX Claim 1; Page 131; 633pp; English.  
 PS  
 CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 396 BP; 100 A; 97 C; 101 G; 98 T; 0 U; 0 Other;  
 Query Match 88.3%; Score 224.4; DB 2; Length 396;  
 Best Local Similarity 97.2%; Pred. No. 3.5e-61;  
 Matches 239; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCAATGCTCAACTTAAGCCGCTTTGTACAGACGATAGACCAACAGGG 60  
 DB 145 ATCAAGCGGACCAATGCTCAACTTAAGCCGCTTTGTACAGACGATAGACCAACAGGG 204  
 QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTTGATCTGGACAGAGAGAG 120  
 DB 205 GAGTTTATGTTTGAATTGATGAAGATGAGAGTTCATGTTGGATCTGGATGAAGAGAG 264  
 QY 121 ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCTTTTCCTTTGAGGCTTCAGGGCGGGCTG 180  
 DB 265 ACCGTCTGGCATCTGGAGGAGTTTGGCCAGAGCTTTTCCTTTGAGGCTTCAGGGCGGGCTG 324  
 QY 181 GCTAACATTGCTATTTGAACAACAACCTTGATACCTTGATCCAGGCTTC-ACCACACT 239  
 DB 325 GCTAACATTGCTATTTGAACAACAACCTTGATACCTTGATCCAGGCTTC-ACCACACT 384  
 QY 240 CAGGCC 245  
 DB 385 CAGGCC 390  
 RESULT 14  
 ID AAV86047 standard; cDNA; 267 BP.  
 AC AAV86047;  
 XX  
 DT 27-APR-1999 (first entry)  
 DE EST clone C180.  
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9845435-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US006954.  
 XX  
 PR 10-APR-1997; 97US-00635913.  
 XX  
 PA (GEM) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

PI Spaulding V, Agostino MJ;  
 XX WPI; 1999-070076/06.  
 DR  
 XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 XX  
 PS Claim 1; Page 101; 633pp; English.  
 CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 267 BP; 69 A; 67 C; 61 G; 61 T; 0 U; 9 Other;  
 Query Match 61.9%; Score 157.2; DB 2; Length 267;  
 Best Local Similarity 90.7%; Pred. No. 7.6e-40;  
 Matches 195; Conservative 0; Mismatches 17; Indels 3; Gaps 3;  
 QY 41 AGACGATAGACCAACAGCGGGAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATG 100  
 DB 1 AGACACATAGACCAACAGCGGGAG-ATATGTTGAATTGATGAAGATGAGATGTTCTATG 59  
 QY 101 TGATCTGGACAAAGAGGAGACCGTCTGGCACTTCGAGGAGTTTGCCAAAGCCTTTCTCT 160  
 DB 60 TGGATCTGACAAAGAGGAGACC-TCTGTCTCTGGAGGAGANTGNCACACCTTTCTCT 118  
 QY 161 TTGAGGCTCAGGGCGGGCTGGCTTAACATTCCTATATTTGACACACAACTTGAATACCTTGA 220  
 DB 119 TTGAGGCTCAGGGCGGGCTGGCTTAACATTCCTATATTTGACACACAACTTGAATACCTTGA 178  
 QY 221 TCCAGGCTTCC-ACCACACTTCAGGCCACCAACCGAT 254  
 DB 179 TCCAGGCTTCCACCACTACACTCAGNCCCAACCAAGAT 213  
 RESULT 15  
 ID AAK54418 standard; cDNA; 267 BP.  
 AC AAK54418;  
 XX  
 DT 13-NOV-2001 (first entry)  
 DE Human haematological malignancy-related antigen coding sequence #143.  
 KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164886-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 01-MAR-2001; 2001WO-US007272.  
 XX  
 PR 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.

```

PR 28-APR-2000; 2000US-0200779P.
PR 01-MAY-2000; 2000US-0200999P.
PR 04-MAY-2000; 2000US-0200804P.
PR 22-MAY-2000; 2000US-0206201P.
PR 14-JUL-2000; 2000US-0218950P.
PR 03-AUG-2000; 2000US-0222903P.
PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 371; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human haematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of haematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Haematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
XX
XX Query Match 57.5%; Score 146; DB 4; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-36;
XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATCAGGCGGACCATGTGTCAACTTATGCCGGTTTGTACAGCGCATAGACCAACAGGG 60
XX 122 ATCAAGGCGGACCATGTGTCAACTTATGCCGGTTTGTACAGCGCATAGACCAACAGGG 181
XX
XX 61 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGCAAGAGAGGAG 120
XX 182 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGCAAGAGAGGAG 241
XX
XX 121 ACCGTCGTCATCTGAGGAGTTGG 146
XX 242 ACCGTCGTCATCTGAGGAGTTGG 267
XX
XX RESULT 16
XX AAK54603
XX ID AAK54603 standard; cDNA; 267 BP.
XX AC AAK54603;
XX
XX DT 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #328.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO2001164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US0007272.

```

```

XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0222903P.
XX 04-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 418; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of haematological malignancies. The
XX present sequence is the coding sequence of a human haematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of haematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Haematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
XX
XX Query Match 57.5%; Score 146; DB 4; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-36;
XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATCAGGCGGACCATGTGTCAACTTATGCCGGTTTGTACAGCGCATAGACCAACAGGG 60
XX 122 ATCAAGGCGGACCATGTGTCAACTTATGCCGGTTTGTACAGCGCATAGACCAACAGGG 181
XX
XX 61 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGCAAGAGAGGAG 120
XX 182 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGCAAGAGAGGAG 241
XX
XX 121 ACCGTCGTCATCTGAGGAGTTGG 146
XX 242 ACCGTCGTCATCTGAGGAGTTGG 267
XX
XX RESULT 17
XX AAK54317
XX ID AAK54317 standard; cDNA; 272 BP.
XX AC AAK54317;
XX
XX DT 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #42.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX

```

PN WO200164886-A2.  
 XX 07-SEP-2001.  
 XX 01-MAR-2001; 2001WO-US007272.  
 XX 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 28-APR-2000; 2000US-0200779P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 22-MAY-2000; 2000US-0202084P.  
 PR 14-JUL-2000; 2000US-0206201P.  
 PR 03-AUG-2000; 2000US-0222903P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX (CORI-) CORIXA CORP.  
 XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 31; Page 345; 1252pp; English.  
 XX The present invention relates to compositions and methods for the detection, diagnosis and therapy of hematological malignancies. The present sequence is the coding sequence of a human hematological malignancy related antigen. The methods of the present invention comprise detecting the presence of hematological malignancy related antigen(s) in a sample obtained from the patient (an increased level of the polypeptide, compared to an unaffected individual, is indicative of an increased risk). Hematological malignancies which can be treated using the present invention are chronic lymphocytic leukemia, lymphoma, follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B cell non-Hodgkin's lymphoma  
 XX Sequence 272 BP; 67 A; 60 C; 75 G; 70 T; 0 U; 0 Other;  
 Query Match 57.5%; Score 146; DB 4; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-36;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATCAAGGGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
 DB 127 ATCAAGGGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 186  
 QY 61 GAGTTTATGTTGAATTGATCAAGATGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 DB 187 GAGTTTATGTTGAATTGATCAAGATGATGTTCTATGTGGATCTGGACAAGAGGAG 246  
 QY 121 ACCGCTCTGGCATCTGGAGGAGTTTGG 146  
 DB 247 ACCGCTCTGGCATCTGGAGGAGTTTGG 272  
 RESULT 18  
 AAK54555/c  
 ID AAK54555 standard; cDNA; 294 BP.  
 XX AAK54555;  
 XX 13-NOV-2001 (first entry)  
 XX Human hematological malignancy-related antigen coding sequence #280.  
 XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;

KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
 XX Homo sapiens.  
 XX WO200164886-A2.  
 XX 07-SEP-2001.  
 XX 01-MAR-2001; 2001WO-US007272.  
 XX 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 04-MAY-2000; 2000US-0202084P.  
 PR 22-MAY-2000; 2000US-0206201P.  
 PR 14-JUL-2000; 2000US-0218950P.  
 PR 03-AUG-2000; 2000US-0222903P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX (CORI-) CORIXA CORP.  
 XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 31; Page 406; 1252pp; English.  
 XX The present invention relates to compositions and methods for the detection, diagnosis and therapy of hematological malignancies. The present sequence is the coding sequence of a human hematological malignancy related antigen. The methods of the present invention comprise detecting the presence of hematological malignancy related antigen(s) in a sample obtained from the patient (an increased level of the polypeptide, compared to an unaffected individual, is indicative of an increased risk). Hematological malignancies which can be treated using the present invention are chronic lymphocytic leukaemia, lymphoma, follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B cell non-Hodgkin's lymphoma  
 XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;  
 Query Match 57.5%; Score 146; DB 4; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-36;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATCAAGGGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
 DB 146 ATCAAGGGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 87  
 QY 61 GAGTTTATGTTGAATTGATCAAGATGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 DB 86 GAGTTTATGTTGAATTGATCAAGATGATGTTCTATGTGGATCTGGACAAGAGGAG 27  
 QY 121 ACCGCTCTGGCATCTGGAGGAGTTTGG 146  
 DB 26 ACCGCTCTGGCATCTGGAGGAGTTTGG 1  
 RESULT 19  
 AAK54805/c  
 ID AAK54805 standard; cDNA; 294 BP.  
 XX AAK54805;  
 XX



DT	13-NOV-2001	(first entry)
XX	Human haematological malignancy-related antigen coding sequence #530.	
DE	Human haematological malignancy-related antigen coding sequence #530.	
XX	Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;	
KW	haematological malignancy; antigen; chronic lymphocytic leukaemia;	
KW	follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200164886-A2.	
XX		
PD	07-SEP-2001.	
XX		
PF	01-MAR-2001; 2001WO-US007272.	
XX		
PR	01-MAR-2000; 2000US-0186136P.	
PR	17-MAR-2000; 2000US-0130479P.	
PR	27-APR-2000; 2000US-0200543P.	
PR	28-APR-2000; 2000US-0200303P.	
PR	28-APR-2000; 2000US-0200779P.	
PR	01-MAY-2000; 2000US-0200999P.	
PR	04-MAY-2000; 2000US-0202084P.	
PR	22-MAY-2000; 2000US-0206201P.	
PR	14-JUL-2000; 2000US-0218950P.	
PR	03-AUG-2000; 2000US-0222903P.	
PR	04-AUG-2000; 2000US-0223416P.	
PR	07-AUG-2000; 2000US-0223378P.	
XX		
PA	(CORI-) CORIXA CORP.	
XX		
XX	Gaiger A, Algate PA, Mannion J;	
PI		
XX	WPI; 2001-514842/56.	
DR		
XX	Compositions and methods for the detection of hematological malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.	
PT		
XX	Claim 31; Page 469; 1252pp; English.	
PS		
XX	The present invention relates to compositions and methods for the detection, diagnosis and therapy of haematological malignancies. The present sequence is the coding sequence of a human haematological malignancy related antigen. The methods of the present invention comprise detecting the presence of haematological malignancy related antigen(s) in a sample obtained from the patient (an increased level of the polypeptide, compared to an unaffected individual, is indicative of an increased risk). Haematological malignancies which can be treated using the present invention are chronic lymphocytic leukaemia, lymphoma, follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B cell non-Hodgkin's lymphoma	
CC		
XX	Sequence 294 BP; 79 A; 77 C; 69 G; 70 T; 0 U; 0 Other;	
XX		
CC	Query March 57.5%; Score 146; DB 4; Length 294;	
CC	Best Local Similarity 100.0%; Pred.No. 2.9e-36;	
CC	Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy	1 ATCAAGGGCGACCTGTCCTCAACTTATCGCGGTTGTACAGACCATAGACCAACAGGG 60	
D5	146 ATCAAGGGCGACCATGTCCTCAACTTATCGCGGTTGTACAGACCATAGACCAACAGGG 87	
Qy	61 CAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGCACAAGAGGAG 120	
D5	86 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGCACAAGAGGAG 27	
Qy	121 ACCGTCGGCATCTCGAGGAGTTTCG 146	
D5	26 ACCGTCGGCATCTCGAGGAGTTTCG 1	
RESULT	20	

Db 561 GCCAACATAGCTGTGGACAAACCAACTTGGAAATCATGACAAAGCGCTCCAAC 614

RESULT 21

AAT97175

ID AAT97175 standard; DNA; 747 BP.

XX AAT97175;

XX 17-OCT-2003 (revised)

DT 11-MAY-1998 (first entry)

XX DR alpha-DAF chimeric gene.

XX Vaccine; B-cell malignancy; lymphoma; leukaemia; tumour;

KW autoimmune disease; gene amplification; immunotherapy; therapy; DAF;

KW decay accelerating factor; DR alpha; mouse; ds.

XX Mus musculus.

OS Chimeric.

PN WO9741244-A1.

XX 06-NOV-1997.

XX 25-APR-1997; 97WO-US007039.

PF 01-MAY-1996; 96US-00644664.

PR 06-DEC-1996; 96US-00761277.

XX (GENI-) GENITOPE CORP.

PA Denney DW;

PI WPI; 1997-549743/50.

DR P-PSDB; AAW37341.

XX Multivalent vaccine to treat B cell lymphoma or leukaemia - comprises at least 2 different recombinant variable regions of immunoglobulin molecules derived from B cell lymphoma cells.

XX Example 5; Page 113-114; 177pp; English.

XX This sequence comprises a chimeric DR alpha-DAF gene, which encodes a chimeric protein (see AAW37341) in which the extracellular domain of mouse DR-alpha is joined to sequences derived from decay accelerating factor (DAF) (see AAT97173). The DAF sequence provides a glycoposphatidylinositol linkage which allows the chimeric protein to be cleaved from the surface of a host cell by treatment of the cell with phospholipase C. The chimeric gene was used in the construction of pSR methods for the expression and co-amplification of genes encoding recombinant proteins in cultured cells. The amplified cells provide large quantities of recombinant proteins suitable for immunotherapy for treatment of lymphomas and leukaemias. The methods permit the production of custom vaccines, including multivalent vaccines, that reflect the degree of somatic variation found in a patient's tumour. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 747 BP; 187 A; 186 C; 187 G; 187 T; 0 U; 0 Other;

SQ

Query Match 45.5%; Score 115.6; DB 2; Length 747;

Best Local Similarity 68.4%; Pred. No. 2.1e-26;

Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCACTGTCTCACTTATCGCGGTTCACAGCCATGACCAACACAGG 60

DB 76 ATCAAGAGAACATGTGATCATCATCAGCGCCGAGTTCTTCTGAATCTGACCAATCAGGC 135

QY 61 GAGTTTATGTGATTTGATGAAGATGATGTTCTATGTGGATCTGGACAAAGAGAG 120

DB 136 GAGTTTATGTGATTTGATGGTATGAGATTTTCCATGTGGATATGCGAAGAGAG 195

QY 121 ACCGCTGGCATCTGGAGGAGTTTGCCAAAGCCCTTTCTTTGAGGCTCAGGCGGGCTG 180

DB 196 ACGGCTGGCGGCTTGAGAAATTTGGACGATTTCCAGGCTTGAAGTGCATTTG 255

QY 181 GCTAACATTTGCTATATTGAACAACAACACTTGAATACCTTGATCCAGCGTTCCACC 234

DB 256 GCCAACATAGCTGTGGACAAACCAACTTGGAAATCATGACAAAGCGCTCCAAC 309

RESULT 22

AAV16866

ID AAV16866 standard; cDNA; 750 BP.

XX AAV16866;

XX 17-AUG-1998 (first entry)

DT DR-alpha extracellular domain-Fos leucine zipper fusion cDNA.

DE Major histocompatibility complex class II; MHC class II; human;

XX fusion protein; HLA-DR2; DRA\*0101; binding domain; Fos;

KW dimerisation domain; allergy; autoimmune disease; vaccine;

KW multiple sclerosis; therapy; ss.

XX Homo sapiens.

OS Synthetic.

OS Chimeric.

XX Key

FT CDS Location/Qualifiers

FT 1..738

FT sig\_peptide /tag= a

FT 1..21

FT /tag= b

FT /note= "3' end of secretory signal"

FT 22..735

FT mat\_peptide /tag= c

FT /product= "DR-alpha extracellular domain-linker-Fos leucine zipper dimerisation domain fusion"

XX WO9806749-A2.

XX 19-FEB-1998.

XX 15-AUG-1997; 97WO-US014503.

PR 16-AUG-1996; 96US-0024077P.

XX (HARD) HARVARD COLLEGE.

XX Wucherpfennig KW, Strominger JL;

WPI; 1998-159459/14.

P-PSDB; AAW46943.

XX New Class II MHC fusion proteins - comprising a MHC Class II binding domain and a dimerisation domain or an immunoglobulin region used for modulating immune responses.

XX Example; Page 55; 76pp; English.

XX This nucleotide sequence codes for a fusion protein (see AAW46943) comprising the extracellular domain of DR-alpha (residues 1-191 of DRA\*0101) and the leucine zipper dimerisation domain of Fos. These components were generated by PCR (see AAV16866-69) using primers designed to include a 7-amino acid linker. The construct was reamplified by PCR to permit cloning into pPIC9 as an in-frame fusion with the alpha-mating factor secretion signal, and expressed in *Pichia pastoris* under control of the AOX1 promoter. A DR-beta chain construct (see AAV16867) was also produced. CHO transfectants assembled and secreted DR alpha-beta heterodimers. The DR2 molecules were shown to specifically bind a human myelin basic protein that is recognised by DR2-restricted T cell clones from multiple sclerosis patients. The invention relates to new soluble monovalent or multivalent Class II MHC fusion proteins comprising a MHC



PN WO200236149-A2.  
 XX  
 PD 10-MAY-2002.  
 XX  
 PF 29-OCT-2001; 2001WO-EP012485.  
 XX  
 PR 30-OCT-2000; 2000EP-00123666.  
 XX  
 PA (UTKU/) UTKU N.  
 XX  
 XX Utku N;  
 PI  
 DR WPI; 2002-508109/54.  
 XX  
 DR P-PSDB; AAG79361.  
 XX  
 XX Use of inhibitors of TIR7 ligand binding for preparing compositions for  
 PT e.g. treating graft versus host disease, autoimmune diseases, allergic  
 PT diseases, infectious diseases, sepsis, and tumors, or for improving wound  
 PT healing.  
 XX  
 XX Claim 3; Fig 1a; 35pp; English.  
 PS  
 CC This sequence encodes the human leukocyte associated antigen (HLA) class  
 CC II alpha chain (HLA DR alpha-chain). The method of the invention allows  
 CC for identification of an inhibitor which interferes with the interaction  
 CC of TIR7 with its ligand (HLA DR alpha-chain) for the preparation of a  
 CC pharmaceutical composition for the treatment of graft versus host  
 CC disease, autoimmune diseases, allergic diseases, infectious diseases,  
 CC sepsis, and tumors, for the improvement of wound healing or for inducing  
 CC or maintaining immune unresponsiveness in a subject. The HLA-class II  
 CC alpha chain or its fragment or derivative is useful for identifying drugs  
 CC for the treatment of an immune disease or tumour  
 XX  
 SQ Sequence 819 BP; 210 A; 200 C; 215 G; 194 T; 0 U; 0 Other;  
 Query Match 45.5%; Score 115.6; DB 6; Length 819;  
 Best Local Similarity 68.4%; Pred. No. 2.2e-26;  
 Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCCATAGACCAACAGGG 60  
 DB 103 ATCAAGAGAACATGTGATCATCCAGCGCGAGTTCTATCTGTAATCTGACCAATCAGGC 162  
 QY 61 GAGTTTATGTTTGAATTTGATCAAGATGAGATGTTCTATGTCGATCTGCACAAAGAGAG 120  
 DB 163 GAGTTTATGTTTGAATTTGATCAAGATGAGATGTTCTATGTCGATCTGCACAAAGAGAG 222  
 QY 121 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180  
 DB 223 ACCGTCCTGGCGGCTTGAAGAAATTTGGACGATTTGGCAGCTTTGAGGCTCAAGGTGATG 282  
 QY 181 GCTAACATGCTATATTGAACAACAACCTTGAATACCTTGCATCCAGCTTCCACC 234  
 DB 283 GCCACATAGCTGTGGACAAAGCCAACTTGGAAATCATGACAAAGGCTCCAC 336

## RESULT 25

ABAB3101  
 ID ABAB3101 standard; DNA; 1183 BP.

AC ABAB3101;

XX 08-FEB-2002 (first entry)

DE HLA-DR alpha chain ovarian tumour marker gene, SEQ ID NO:40.

XX Ovarian tumour marker gene; human; overexpression; upregulation;  
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;  
 KW identification; serous cystadenoma; borderline serous tumour;  
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;  
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;  
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;  
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;

KW immune response pathway; cell proliferation regulation; protein folding;  
 KW membrane localised; secreted; therapeutic target; cytostatic;  
 KW gene therapy; vaccine; ds.  
 XX Homo sapiens.  
 OS  
 XX WO200175177-A2.  
 PN  
 XX 11-OCT-2001.  
 PD  
 XX 03-APR-2001; 2001WO-US010947.  
 XX  
 XX 03-APR-2000; 2000US-0194336P.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;  
 PI  
 XX WPI; 2001-626450/72.  
 XX  
 XX P-PSDB; ABB50276.  
 DR  
 XX  
 XX Detecting and identifying ovarian tumor, identifying increased risk for  
 PT developing ovarian cancer, and determining effectiveness of ovarian  
 PT cancer treatment, by measuring expression level of ovarian tumor marker  
 PT gene.  
 XX  
 PS Claim 23; Page 89; 140pp; English.  
 XX  
 CC The invention relates to methods for diagnosing and prognosing ovarian  
 CC tumors in an individual via the detection and measurement of the  
 CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,  
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,  
 CC ABA83181 and ABA83183). The methods of the invention are useful for  
 CC detecting an ovarian tumour in a patient, for identifying an individual  
 CC at increased risk for developing ovarian cancer, in prognostic tests for  
 CC assessing the relative severity of ovarian cancer, in tests for  
 CC monitoring a patient in remission from ovarian cancer and in tests for  
 CC monitoring disease status in a patient being treated for ovarian cancer.  
 CC The methods can additionally be used to identify a particular tumour as  
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from  
 CC mucinous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,  
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous  
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,  
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner  
 CC tumour. The ovarian tumour marker genes of the invention were identified  
 CC using SAGE (serial analysis of gene expression) and were found to be  
 CC overexpressed in a broad variety of ovarian epithelial tumour cells  
 CC relative to normal ovarian epithelial cells. The marker genes are  
 CC implicated in immune response pathways, in the regulation of cell  
 CC proliferation and in protein folding, and many of these are membrane-  
 CC localised or secreted. In addition to their use as diagnostic and  
 CC prognostic markers, the ovarian tumour marker genes or their encoded  
 CC proteins may be used as therapeutic targets for the treatment and  
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,  
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of the  
 CC invention

SQ Sequence 1183 BP; 295 A; 286 C; 278 G; 321 T; 0 U; 1 Other;

Query Match 45.5%; Score 115.6; DB 5; Length 1183;

Best Local Similarity 68.4%; Pred. No. 2.5e-26;

Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCCATAGACCAACAGGG 60  
 DB 90 ATCAAGAGAACATGTGATCATCCAGCGCGAGTTCTATCTGTAATCTGACCAATCAGGC 149  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTCGATCTGCACAAAGAGAG 120  
 DB 150 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTCGATCTGCACAAAGAGAG 209  
 QY 121 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180

```

Db      210  ACGGCTCGGGCTTTGAAGAAATTGGACGATTTGCCAGCTTTGCGGCTCAAGGTGCATTG 269
QY      -81  GCTACATTCGTATATTGAACAACAACTTGAATACCTTGATCCAGCGTTCACCC 234
        ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db      270  GCCAACATAGCTGTGGACAAAGCCCAACTTGGAAATCATGACAAAGCGCTCCAAC 323

RESULT 26
AAT99707
XX      ID  AAT99707 standard; cDNA; 1446 BP.
XX      AC  AAT99707;
XX      DT  17-OCT-2003 (revised)
XX      DT  17-AUG-1998 (first entry)
XX      DE  DR2-IgG fusion construct.
XX      DE  Major histocompatibility complex class II; MHC class II; human; mouse;
XX      KW  fusion protein; HLA-DR2; DR2*0101; binding domain; Fos; vaccine;
XX      KW  dimerisation domain; IgG; allergy; autoimmune disease;
XX      KW  multiple sclerosis; therapy; ss.
XX      OS  Homo sapiens.
XX      OS  Mus musculus.
XX      OS  Chimeric.
XX      FN  WO9806749-A2.
XX      PD  19-FEB-1998.
XX      FF  15-AUG-1997; 97WO-US014503.
XX      PR  16-AUG-1996; 96US-0024077P.
XX      FA  (HARD ) HARVARD COLLEGE.
XX      PI  Wucherpennig KW, Strominger JL;
XX      XX  WPI; 1998-159459/14.
XX      XX  New Class II MHC fusion proteins - comprising a MHC Class II binding
XX      PT  domain and a dimerisation domain or an immunoglobulin region used for
XX      PT  modulating immune responses.
XX      PS  Example; Page 49; 76pp; English.
XX      CC  This nucleotide sequences codes for a bivalent DR2 fusion protein
XX      CC  obtained by fusion of the Fc portion of IgG2a to the 3' end of a DR-alpha
XX      CC  -Fos cDNA construct (see AAV15866). The Fc portion was amplified by RT-
XX      CC  PCR from mouse hybridoma L243. The PCR product was then fused in frame
XX      CC  with the DR-alpha-Fos construct by overlapping PCR. The DR2-IgG fusion
XX      CC  was expressed in the Drosophila Schneider cell system. The invention
XX      CC  relates to new soluble monovalent and multivalent Class II MHC fusion
XX      CC  proteins comprising a MHC Class II binding domain and a dimerisation
XX      CC  domain or an immunoglobulin region that can be used for the treatment of
XX      CC  allergic and autoimmune diseases (e.g. multiple sclerosis), for
XX      CC  tolerising a subject to foreign tissue before or after organ or tissue
XX      CC  transplantation, or for vaccination against pathogens. (Updated on 17-OCT
XX      CC  -2003 to standardise OS field)
XX      SQ  Sequence 1446 BP; 414 A; 375 C; 356 G; 301 T; 0 U; 0 Other;

Query Match      45.5%; Score 115.6; D3 2; Length 1446;
Best Local Similarity 68.4%; Pred. No. 2.7e-26;
Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY      1  ATCAAGCGGACCATGTGTCAACTTATCGCGCTTTGTACAGACCATAGACCAACAGCG 60
        ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db      16  ATCAAGAGACATGTATCATCCAGCGCGAGTCTATCTGAATCTTCACCAATCAGGC 75
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      61  GAGTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAAGAGGAG 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      76  GAGTTTATGTTGACTTTGATGGTGATGAGATTTTCCATGTGGATATGGCAAGAGAG 135
QY      121  ACCGCTCGGATCTGGAGGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGCTG 180
        ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db      136  ACGTCTCGGGCTTGAAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAAGTGCATTG 195
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      181  GCTAACATTCGTATATTGAACAACAACTTGAATACCTTGATCCAGCGTTCACCC 234
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      196  GCCAACATAGCTGTGGACAAAGCCCAACTTGGAAATCATGACAAAGCGCTCCAAC 249

RESULT 27
AAX87813
XX      ID  AAX87813 standard; DNA; 1446 BP.
XX      AC  AAX87813;
XX      DT  09-NOV-1999 (first entry)
XX      DE  HLA-DR2 alpha-Fos-IgG fusion construct.
XX      DE  Major histocompatibility complex Class II; MHC; binding domain; HLA-DR2;
XX      KW  leucine zipper; Fos; IgG; Fc; immunoglobulin; antibody; fusion protein;
XX      KW  multiple sclerosis; rheumatoid arthritis; graft rejection; allergy;
XX      KW  autoimmune disease; pemphigus vulgaris; systemic lupus erythematosus;
XX      KW  T lymphocyte; T cell; diagnosis; therapy; adoptive immunotherapy; ss.
XX      OS  Homo sapiens.
XX      OS  Saccharomyces cerevisiae.
XX      OS  Synthetic.
XX      OS  Chimeric.
XX      FH  Key      Location/Qualifiers
XX      FT  CDS      1..1440
XX      FT          /*tag= a
XX      FT          1..15
XX      FT          /*tag= b
XX      FT          /note= "alpha-mating factor secretion signal"
XX      FT  mat_peptide 16..1437
XX      FT          /*tag= c
XX      FT          /product= "DR2-Fos-Fc"
XX      PN  WO9942597-A1.
XX      PD  26-AUG-1999.
XX      PP  19-FEB-1999; 99WO-US003603.
XX      PR  19-FEB-1998; 98US-0075351P.
XX      PA  (HARD ) HARVARD COLLEGE.
XX      PI  Wucherpennig KW, Strominger JL;
XX      XX  WPI; 1999-527481/44.
XX      DR  P-PSDB; AAY31654.
XX      DR  P-PSDB; AAY31654.
XX      PT  New HMC Class II binding domain fusion proteins and conjugates - used
XX      PT  for, e.g. treating allergic and autoimmune diseases or detecting,
XX      PT  isolating, activating or killing specific T cells.
XX      PS  Example 7; Page 100-102; 113pp; English.
XX      CC  This nucleotide sequence codes for a divalent HLA-DR2 MHC binding domain
XX      CC  fusion protein (see AAY31654) comprising an alpha-mating factor secretion
XX      CC  signal, the extracellular domain of the HLA-DR2 alpha chain (residues 1-
XX      CC  191 of DR2*0101), a 7-amino acid linker, the 40-amino acid leucine zipper
XX      CC  dimerization domain of Fos, and the Fc portion of IgG2a. The DR-alpha-Fc
XX      CC  chain corresponds to an antibody heavy chain. The invention provides new
XX      CC  monovalent, multivalent and multimeric MHC Class II binding domain fusion
XX      CC  proteins and conjugates comprising at least a binding domain of an MHC
XX      CC  Class II alpha or beta chain and a dimerization domain, especially a Fos
XX      CC  or Jun leucine zipper domain. The MHC fusion proteins and conjugates can

```

CC be used; for detecting and isolating T cells having a defined MHC/peptide  
 CC complex specificity (claimed); to confer to a subject adoptive immunity  
 CC to a defined MHC/peptide complex (claimed); to stimulate or activate T  
 CC cells reactive to a defined MHC/peptide complex (claimed); for selective  
 CC killing of T cells reactive to a defined MHC/peptide complex (claimed); to  
 CC tolerate a subject to a defined MHC/peptide complex (claimed); to treat  
 CC allergic and autoimmune diseases, e.g. multiple sclerosis, rheumatoid  
 CC arthritis, pemphigus vulgaris, and systemic lupus erythematosus; and to  
 CC prevent organ or tissue transplant rejection. The DR2-IGG design was  
 CC chosen to increase the affinity for the T cell receptor by increasing  
 CC valency, and to attach an effector domain, the Fc region of IgG2a.  
 CC Complement fixation may result in the lysis of target T cells following  
 CC binding of DR2-IGG molecules to the T cell receptor. DR2-IGG molecules  
 CC may therefore be useful for the selective depletion of autoaggressive T  
 CC cells

XX Sequence 1446 BP; 414 A; 375 C; 356 G; 301 T; 0 U; 0 Other;  
 Query Match 45.5%; Score 115.6; DB 2; Length 1446;  
 Best Local Similarity 68.4%; Pred. No. 2.7e-26;  
 Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCACTGTCACACTATGCGCGGTTGTACAGACCATAGACCAACAGG 60  
 Db 16 ATCAAGAAGAACATGTATCATCCAGGCGGAGTCTATCTGAATCCTGACCAATCAGG 75  
 QY 61 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGTGATCTCGACAAGAGGAG 120  
 Db 76 GAGTTTATGTTGACTTTGAATGGTATGAGATTTTCCATGTGATATGGCAAGAGGAG 135  
 QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGGCAAGCCTTTCTTTGAGGCTTCAGGCGGGCTG 180  
 Db 136 ACGETCTGGCGGCTTGAAGAAATTTGGACCATTTGGCAGCTTTGAGGCTCAAGTGCATG 195  
 QY 181 GCTACATGCTATATTGAACACAACTGATATCTGATACCTTGATCCAGCCTTCACC 234  
 Db 196 GCCAACATAGCTGGACAAAGCCAACTTGGAAATCATGACAAAGCGCTCCAAC 249

RESULT 28  
 AAT99708  
 ID AAT99708 standard; cDNA; 1851 BP.  
 XX  
 AC AAT99708;

XX 17-OCT-2003 (revised)  
 DT 17-AUG-1998 (first entry)  
 XX  
 DE DR2-IGM fusion construct.  
 XX Major histocompatibility complex class II; MHC class II; human; mouse;  
 KW fusion protein; HLA-DR2; DRA\*0101; binding domain; Fos;  
 KW dimerisation domain; IGM; allergy; autoimmune disease; vaccine;  
 XX multiple sclerosis; therapy; ss.

XX Homo sapiens.  
 OS Mus musculus.  
 OS Chimeric.  
 XX  
 PN W09806749-A2.  
 XX  
 XX 19-FEB-1998:  
 XX  
 PF 15-AUG-1997; 97WO-US014503.  
 XX  
 PR 16-AUG-1996; 96US-0024077P.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Wucherpfennig KW, Strominger JL;  
 DR WPI; 1998-159459/14.  
 XX

PT New Class II MHC fusion proteins - comprising a MHC Class II binding  
 PT domain and a dimerisation domain or an immunoglobulin region used for  
 XX modulating immune responses.

XX Example; Page 49-50; 76pp; English.

XX This nucleotide sequences codes for a DR2 fusion protein obtained by  
 CC fusion of the Fc portion of murine IgM to the 3' end of a DR-alpha-Fos  
 CC cDNA construct (see AAV16866). DR2-IGM molecules carry 10 DR2/peptide  
 CC arms and are expected to have a higher functional affinity for the T cell  
 CC receptor than a DR2-IGG fusion (see AAT99707). DR2-IGM fusion molecules  
 CC were secreted from transfected COS cells. The fusion protein may be of  
 CC utility for the depletion of antigen-specific T cells in patients with  
 CC multiple sclerosis. The invention relates to new soluble monovalent and  
 CC multivalent Class II MHC fusion proteins comprising a MHC Class II  
 CC binding domain and a dimerisation domain or an immunoglobulin region that  
 CC can be used for the treatment of allergic and autoimmune diseases, for  
 CC tolerising a subject to foreign tissue before or after organ or tissue  
 CC transplantation, or for vaccination against pathogens. (Updated on 17-OCT  
 CC -2003 to standardise OS field)

XX Sequence 1851 BP; 485 A; 503 C; 447 G; 416 T; 0 U; 0 Other;  
 Query Match 45.5%; Score 115.6; DB 2; Length 1851;  
 Best Local Similarity 68.4%; Pred. No. 3e-26;  
 Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCACTGTCACACTATGCGCGGTTGTACAGACCATAGACCAACAGG 60  
 Db 76 ATCAAGAAGAACATGTATCATCCAGGCGGAGTCTATCTGAATCCTGACCAATCAGG 135  
 QY 61 GAGTTTATGTTGAAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120  
 Db 136 GAGTTTATGTTGACTTTGAATGGTATGAGATTTTCCATGTGATATGGCAAGAGGAG 195  
 QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGGCAAGCCTTTCTTTGAGGCTTCAGGCGGGCTG 180  
 Db 196 ACGETCTGGCGGCTTGAAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAAGTGCATG 255  
 QY 181 GCTACATGCTATATTGAACACAACTGATATCTGATACCTTGATCCAGCCTTCACC 234  
 Db 256 GCCAACATAGCTGGACAAAGCCAACTTGGAAATCATGACAAAGCGCTCCAAC 309

RESULT 29  
 AAX87814  
 ID AAX87814 standard; DNA; 1851 BP.  
 XX  
 AC AAX87814;  
 XX  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE HLA-DR2 alpha-Fos-IgM fusion construct.

XX Major histocompatibility complex Class II; MHC; binding domain; HLA-DR2;  
 KW leucine zipper; Fos; IGM; FC; immunoglobulin; antibody; fusion protein;  
 KW multiple sclerosis; rheumatoid arthritis; graft rejection; allergy;  
 KW autoimmune disease; pemphigus vulgaris; systemic lupus erythematosus;  
 KW T lymphocyte; T cell; diagnosis; therapy; adoptive immunotherapy; ss.

XX Homo sapiens.  
 OS Saccharomyces cerevisiae.  
 OS Synthetic.  
 OS Chimeric.

XX  
 EH Key Location/Qualifiers  
 FT CDS 1..1839  
 FT /\*tag= a  
 FT sig\_peptide 1..75  
 FT /\*tag= b  
 FT /note= "alpha-mating factor secretion signal"  
 FT mat\_peptide 76..1836  
 FT /\*tag= c

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 12:38:49 ; Search time 56 Seconds

(without alignments)

2517.099 Million cell updates/sec

Title: US-09-877-819B-55

Perfect score: 254

Sequence: 1 atcaaggcgaccatgtgc.....acactcaggccaccacogat 254

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 692709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	115.6	45.5	747	1	US-08-644-664B-26
2	115.6	45.5	747	2	US-08-761-277A-26
3	114	44.9	690	3	US-08-470-535-11
4	114	44.9	1244	3	US-08-463-903-23
5	114	44.9	1244	3	US-08-463-903-25
6	114	44.9	1244	3	US-08-463-903-27
7	114	44.9	1244	3	US-08-463-903-29
8	114	44.9	1244	4	US-07-935-695-23
9	114	44.9	1244	4	US-07-935-695-25
10	114	44.9	1244	4	US-07-935-695-27
11	114	44.9	1244	4	US-07-935-695-29
12	92.6	36.5	776	3	US-08-462-351-4
13	92.6	36.5	776	4	US-03-602-807-4
14	92.6	36.5	776	6	US1942425-5
15	91	35.8	776	6	5468481-4
16	87.6	34.5	1382	2	US-08-596-387B-123
17	87.6	34.5	1382	4	US-09-067-615-123
18	87.6	34.5	1382	5	PCT-US95-09816A-123
19	87.6	34.5	1382	2	US-08-596-387B-121
20	87.6	34.5	1382	4	US-09-067-615-121
21	87.6	34.5	1382	5	PCT-US95-09816A-121
22	87.6	34.5	1508	2	US-08-596-387B-122
23	87.6	34.5	1508	3	US-08-960-190A-24
24	87.6	34.5	1508	4	US-09-067-615-22
25	87.6	34.5	1508	5	PCT-US95-09816A-122
26	87.6	34.5	4713	4	US-09-194-285-7
27	83.4	32.8	956	4	US-09-620-312D-229

3	US-09-153-586-1	566	32.6	82.8	28	Sequence 1, Appli
4	US-08-332-616A-4	242	32.2	81.8	29	Sequence 4, Appli
5	US-08-317-220-4	242	32.2	81.8	30	Sequence 4, Appli
6	5451505-3	188	31.8	80.8	31	Patent No. 5451505
7	US-07-775-786-25	57	22.0	56	32	Sequence 25, Appl
8	US-07-775-786-26	57	22.0	56	33	Sequence 26, Appl
9	US-09-258-133-23	57	22.0	56	34	Sequence 23, Appl
10	US-07-775-786-23	57	21.7	55	35	Sequence 23, Appl
11	US-09-258-133-22	57	21.7	55	36	Sequence 22, Appl
12	US-07-775-786-24	57	21.0	53.4	37	Sequence 24, Appl
13	US-08-070-328A-1	150	17.9	45.4	38	Sequence 1, Appli
14	US-08-455-116-1	150	17.9	45.4	39	Sequence 1, Appli
15	US-08-936-165A-119	1406	13.3	33.8	40	Sequence 119, App
16	US-08-956-171B-3054	382	12.8	32.5	41	Sequence 3054, Ap
17	US-08-956-171B-2979	400	12.6	32	42	Sequence 2379, Ap
18	US-08-356-171B-444	11466	12.6	32	43	Sequence 444, App
19	US-09-535-008-52	1513	12.3	31.2	44	Sequence 52, Appl
20	US-09-621-976-8976	399	12.0	30.6	45	Sequence 8976, Ap
21	US-09-465-558-29	571	12.0	30.4	46	Sequence 29, Appl
22	US-08-916-421B-1	1664976	12.0	30.4	47	Sequence 1, Appli
23	US-09-103-840A-2	4403765	11.9	30.2	48	Sequence 2, Appli
24	US-09-103-840A-1	4411529	11.9	30.2	49	Sequence 88, Appli
25	US-08-318-193-88	412	11.8	30	50	Sequence 107, App
26	US-09-227-357-107	1146	11.7	29.8	51	Sequence 60, Appl
27	US-09-535-008-60	5386	11.7	29.6	52	Sequence 66, Appl
28	US-09-535-008-66	5468	11.7	29.6	53	Sequence 1, Appli
29	US-09-535-008-1	5471	11.7	29.6	54	Sequence 62, Appl
30	US-09-535-008-62	5471	11.7	29.6	55	Sequence 74, Appl
31	US-09-535-008-74	5477	11.7	29.6	56	Sequence 70, Appl
32	US-09-535-008-70	5480	11.7	29.6	57	Sequence 68, Appl
33	US-09-535-008-68	5564	11.7	29.6	58	Sequence 64, Appl
34	US-09-535-008-64	5567	11.7	29.6	59	Sequence 76, Appl
35	US-09-535-008-76	5573	11.7	29.6	60	Sequence 72, Appl
36	US-09-535-008-72	5576	11.7	29.6	61	Sequence 9, Appli
37	US-09-167-206-9	1111	11.6	29.4	62	Sequence 2, Appli
38	US-09-103-840A-2	4403765	11.6	29.4	63	Sequence 1, Appli
39	US-09-103-840A-1	4411529	11.6	29.4	64	Sequence 2, Appli
40	US-08-431-560-2	2499	11.5	29.2	65	Sequence 2, Appli
41	US-08-463-345-2	2499	11.5	29.2	66	Sequence 1, Appli
42	US-09-412-545-1	5032	11.5	29.2	67	Sequence 16380, A
43	US-09-621-976-16380	400	11.4	29	68	Sequence 2, Appli
44	US-08-766-738-2	966	11.4	29	69	Sequence 2, Appli
45	US-09-252-610-2	966	11.4	29	70	Sequence 5, Appli
46	US-09-470-443-5	5279	11.4	29	71	Sequence 1, Appli
47	US-09-470-443-1	5463	11.4	29	72	Sequence 3, Appli
48	US-09-470-443-3	5482	11.4	29	73	Sequence 385, App
49	US-09-833-381-985	944	11.3	28.8	74	Sequence 51, Appl
50	US-09-566-921-51	2333	11.3	28.8	75	Sequence 1, Appli
51	US-07-945-283-1	8438	11.3	28.8	76	Sequence 16, Appl
52	US-09-442-099A-16	196	11.3	28.6	77	Sequence 16, Appl
53	US-09-612-342-16	196	11.3	28.6	78	Sequence 15, Appl
54	US-09-612-421A-16	196	11.3	28.6	79	Sequence 15, Appl
55	US-09-252-991A-13549	369	11.3	28.6	80	Sequence 13549, A
56	US-09-276-531-60	3436	11.3	28.6	81	Sequence 60, Appl
57	US-09-385-982-222	617	11.2	28.4	82	Sequence 222, App
58	US-09-043-123-1	1557	11.2	28.4	83	Sequence 1, Appli
59	US-09-134-000C-2821	1560	11.2	28.4	84	Sequence 2821, Ap
60	US-08-454-557C-50	2520	11.2	28.4	85	Sequence 50, Appl
61	US-08-340-426D-50	2520	11.2	28.4	86	Sequence 50, Appl
62	PCT-US95-1711A-50	2520	11.2	28.4	87	Sequence 50, Appl
63	US-09-368-590-1	7812	11.2	28.4	88	Sequence 1, Appli
64	US-08-961-527-164	9707	11.2	28.4	89	Sequence 164, App
65	US-09-134-000C-1921	1542	11.1	28.2	90	Sequence 1921, Ap
66	US-09-620-312D-492	4932	11.1	28.2	91	Sequence 492, App
67	US-09-620-312D-491	4964	11.1	28.2	92	Sequence 491, App
68	US-08-207-481-44	215	11.0	28	93	Sequence 44, Appl
69	PCT-US95-02689-51	215	11.0	28	94	Sequence 51, Appl
70	US-08-748-506-9	984	11.0	28	95	Sequence 9, Appli
71	US-08-207-481-38	1013	11.0	28	96	Sequence 38, Appl
72	PCT-US95-02689-40	1013	11.0	28	97	Sequence 40, Appl
73	US-09-205-258-210	2110	11.0	28	98	Sequence 210, App
74	US-09-134-000C-1956	3162	11.0	28	99	Sequence 1956, Ap
75		100			100	

```

c 101 28 11.0 21338 4 US-08-961-527-20 Sequence 20, Appl
c 102 27.8 10.9 535 4 US-09-569-751-123 Sequence 123, Appl
c 103 27.8 10.9 1091 4 US-09-565-655-62 Sequence 62, Appl
c 104 27.8 10.9 2141 4 US-09-441-039-1 Sequence 1, Appl
c 105 27.8 10.9 2142 4 US-09-441-039-3 Sequence 3, Appl
c 106 27.8 10.9 2149 1 US-07-623-953-2 Sequence 2, Appl
c 107 27.8 10.9 3093 4 US-09-517-529A-1 Sequence 1, Appl
c 108 27.8 10.9 9301 4 US-09-449-218D-18 Sequence 18, Appl
c 109 27.8 10.9 9301 4 US-09-668-529A-18 Sequence 18, Appl
c 110 27.8 10.9 9301 4 US-09-668-037A-12 Sequence 12, Appl
c 111 27.8 10.9 92139 4 US-09-518-686-1 Sequence 1, Appl
c 112 27.6 10.9 577 4 US-09-621-976-2981 Sequence 2981, Ap
c 113 27.4 10.8 397 4 US-09-401-064-232 Sequence 232, App
c 114 27.4 10.8 397 4 US-09-401-064-244 Sequence 244, App
c 115 27.4 10.8 1042 4 US-09-072-596-331 Sequence 331, App
c 116 27.4 10.8 1042 4 US-09-072-967-336 Sequence 336, App
c 117 27.4 10.8 1389 4 US-09-489-039A-6445 Sequence 6445, Ap
c 118 27.4 10.8 1355 4 US-09-489-039A-6966 Sequence 6966, Ap
c 119 27.4 10.8 14507 3 US-08-785-150-1 Sequence 1, Appl
c 120 27.4 10.8 14507 4 US-09-660-299-1 Sequence 1, Appl
c 121 27.4 10.8 14507 4 US-09-435-377-1 Sequence 1, Appl
c 122 27.4 10.8 14507 4 US-09-973-928-1 Sequence 1, Appl
c 123 27.4 10.8 161852 4 US-09-497-858A-40 Sequence 40, Appl
c 124 27.2 10.7 1354 2 US-08-587-680A-24 Sequence 24, Appl
c 125 27.2 10.7 1559 2 US-08-943-087-13 Sequence 13, Appl
c 126 27.2 10.7 1659 2 US-08-943-087-21 Sequence 21, Appl
c 127 27.2 10.7 1659 2 US-08-943-087-23 Sequence 23, Appl
c 128 27.2 10.7 1659 2 US-08-943-087-41 Sequence 41, Appl
c 129 27.2 10.7 1659 2 US-08-943-087-43 Sequence 43, Appl
c 130 27.2 10.7 1668 4 US-09-620-312D-127 Sequence 127, App
c 131 27.2 10.7 2271 1 US-09-904-615-52 Sequence 52, Appl
c 132 27.2 10.7 2338 1 US-08-425-069-1 Sequence 1, Appl
c 133 27.2 10.7 2338 2 US-08-317-844B-1 Sequence 1, Appl
c 134 27 10.6 486 4 US-09-252-991A-14270 Sequence 14270, A
c 135 27 10.6 612 3 US-09-385-982-566 Sequence 566, App
c 136 27 10.6 1020 4 US-09-328-111-172 Sequence 172, App
c 137 27 10.6 16568 4 US-09-252-991A-14330 Sequence 14330, A
c 138 27 10.6 16568 4 US-09-525-906-1 Sequence 1, Appl
c 139 27 10.6 16569 3 US-09-097-889-2 Sequence 2, Appl
c 140 27 10.6 16569 4 US-09-377-856-1 Sequence 2, Appl
c 141 27 10.6 16569 4 US-09-302-681-2 Sequence 2, Appl
c 142 27 10.6 18569 4 US-09-098-079-2 Sequence 2, Appl
c 143 27 10.6 18569 4 US-09-621-976-18033 Sequence 18033, A
c 144 26.8 10.6 2324 4 US-09-508-824-1 Sequence 1, Appl
c 145 26.8 10.6 174493 4 US-09-804-471A-3 Sequence 3, Appl
c 146 26.8 10.6 174493 4 US-10-238-709-3 Sequence 3, Appl
c 147 26.6 10.5 610 4 US-09-612-473-39 Sequence 39, Appl
c 148 26.6 10.5 903 4 US-09-107-532A-1970 Sequence 1970, Ap
c 149 26.6 10.5 964 4 US-09-023-685-1081 Sequence 2, Appl
c 150 26.6 10.5 3387 4 US-09-194-640A-2 Sequence 2, Appl
c 151 26.6 10.5 3423 4 US-09-620-312D-126 Sequence 126, App
c 152 26.6 10.5 3711 4 US-09-566-921-126 Sequence 126, App
c 153 26.6 10.5 3868 3 US-09-357-070-1 Sequence 1, Appl
c 154 26.6 10.5 5053 4 US-09-378-330-1 Sequence 1, Appl
c 155 26.6 10.5 5220 2 US-08-777-405A-1 Sequence 1, Appl
c 156 26.6 10.5 5220 2 US-08-977-871A-1 Sequence 1, Appl
c 157 26.6 10.5 5220 2 US-09-225-951-1 Sequence 1, Appl
c 158 26.6 10.5 5220 4 US-09-841-341-1 Sequence 1, Appl
c 159 26.6 10.5 5220 4 US-10-027-581-1 Sequence 1, Appl
c 160 26.6 10.5 81001 4 US-09-750-580-1 Sequence 1, Appl
c 161 26.6 10.4 300 4 US-09-313-294A-4024 Sequence 4024, Ap
c 162 26.4 10.4 978 3 US-09-237-543-3 Sequence 3, Appl
c 163 26.4 10.4 978 4 US-09-644-450-3 Sequence 3, Appl
c 164 26.4 10.4 1037 3 US-09-116-498-3 Sequence 3, Appl
c 165 26.4 10.4 1288 1 US-08-143-897-4 Sequence 4, Appl
c 166 26.4 10.4 1422 4 US-09-489-039A-3358 Sequence 3358, Ap
c 167 26.4 10.4 1431 4 US-09-465-538-67 Sequence 67, Appl
c 168 26.4 10.4 1440 4 US-09-489-039A-5323 Sequence 5323, Ap
c 169 26.4 10.4 1443 4 US-09-252-991A-12097 Sequence 12097, A
c 170 26.4 10.4 1493 2 US-08-752-307B-6 Sequence 6, Appl
c 171 26.4 10.4 1493 4 US-09-707-802-6 Sequence 6, Appl
c 172 26.4 10.4 1493 4 US-09-991-326-6 Sequence 6, Appl
c 173 26.4 10.4 1493 4 US-09-991-326-6 Sequence 6, Appl

174 26.4 10.4 1551 4 US-09-252-991A-12323 Sequence 12323, A
c 175 26.4 10.4 2085 4 US-09-252-991A-12337 Sequence 12337, A
c 176 26.4 10.4 2126 3 US-09-237-543-1 Sequence 1, Appl
c 177 26.4 10.4 2126 4 US-09-644-450-1 Sequence 1, Appl
c 178 26.4 10.4 2139 4 US-09-489-039A-3565 Sequence 3565, Ap
c 179 26.4 10.4 3747 2 US-09-080-897-1 Sequence 1, Appl
c 180 26.4 10.4 3747 3 US-09-323-735-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-644-664B-26
; Sequence 26, Application US/08644664B
; Patent No. 5776746
; GENERAL INFORMATION:
; APPLICANT: Denney Jr., Dan W.
; TITLE OF INVENTION: Gene Amplification Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/644,664B
; FILING DATE: 01-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: GENITOPPE-00912
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..744
; US-08-644-664B-26

Query Match 45.5%; Score 115.6; DB 1; Length 747;
Best Local Similarity 68.4%; Pred. No. 2.4e-30;
Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 1 ATCAAGCGGACCATGTCTCACTTATGCGCGTGTTCACAGAGCATAGACCAACAGGG 60
Db 76 ATCAAGAGAACATGTGATCATCCAGGCGAGTCTATCTGATCTGTACCATCAGGC 135
Qy 61 GAGTTTATGTTGATTTGATGAGATGAGATGTTCTATGTTGATCTGTCGACAGAGGAG 120
Db 136 GAGTTTATGTTGATTTGATGAGATGAGATTTTCCATGTCGATGATGCAAGAGGAG 195
Qy 121 ACCCTCGGCATCTGGAGGAGTTTGCCAGACCTTTTCCTTTCAGACTCAGGCGGCTG 180
Db 196 ACGTCTGCGCGCTTGAGGATTTTGACGATTTTGCCAGCTTTGAGGCTCAAGTGTGATTG 255
Qy 181 GCTAACATCTCTATTTGAAACAACACTTGAATACCTTGAATCCAGGCTTCCACC 234

```



Ds 256 GCCAACATAGCTGTGGACAAACCAACTTGGAAATCATGACAAAGCGCTCCAAC 309

## RESULT 2

US-08-761-277A-26

; Sequence 26, Application US/08761277A

; Patent No. 5972334

; GENERAL INFORMATION:

; APPLICANT: Denney Jr., Dan W.

; TITLE OF INVENTION: Vaccines For Treatment Of Lymphoma And

; TITLE OF INVENTION: Leukemia

; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States Of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/761,277A

; FILING DATE: 06-DEC-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/644,664

; FILING DATE: 01-MAY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: MacKnight, Karin T.

; REGISTRATION NUMBER: 38,230

; REFERENCE/DOCKET NUMBER: GENITOPE-02406

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 747 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..744

US-08-761-277A-26

Query Match 45.5%; Score 115.6; DB 2; Length 747;

Best Local Similarity 68.4%; Pred. No. 2.4e-30;

Matches 160; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 1 ATCAAGCGGACCATGTGTCACCTTATGCCCGGTTGTACAGACGATAGCCAAACAGGG 60

Ds 76 ATCAAGAGAGAACATGTGATCATCCAGGCCGAGTTCTATCTGAATCTCTGACCAATCAGGC 135

Qy 61 GAGTTTATGTTGATTTGATGAGATGAGATGTTCTATGTTGGATCTGGACAGAGAG 120

Ds 136 GAGTTTATGTTGACTTTGATGGTGATGAGATTTCCATGTTGGATATGGCAAGAGAG 195

Qy 121 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCTTTGAGGCTCAGGGGGGCTG 180

Ds 196 ACGGTCCTGGCGCTTGAGATTTGGAGATTTCCAGCATTTGCCAGCTTTGAGGCTCAGGTCATG 255

Qy 181 GCTAACATGCTATATTGAAACAAACACTTGAATACCTTGTATCCAGGCTCCACC 234

Ds 256 GCCAACATAGCTGTGGACAAACCAACTTGGAAATCATGACAAAGCGCTCCAAC 309

## RESULT 3

US-08-470-535-11

; Sequence 11, Application US/08470535

; Patent No. 6090587

; GENERAL INFORMATION:

; APPLICANT: Rhodes, Eric T

; APPLICANT: Nag, Bishwajit

; TITLE OF INVENTION: PROKARYOTIC EXPRESSION OF MHC PROTEINS

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew

; STREET: One Market Plaza, Steuart Street Tower

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94105-1492

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/470,535

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/143,575

; FILING DATE: 25-OCT-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/329,010

; FILING DATE: 25-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 14058-21-1-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 690 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..690

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 577..690

; OTHER INFORMATION: /note= "Encodes the transmembrane

; OTHER INFORMATION: region within HLA DR2-Dw2 Alpha Chain."

US-08-470-535-11

Query Match 44.9%; Score 114; DB 3; Length 690;

Best Local Similarity 67.9%; Pred. No. 8.4e-30;

Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 1 ATCAAGCGGACCATGTGTCACCTTATGCCCGGTTGTACAGACGATAGCCAAACAGGG 60

Ds 1 ATCAAGAGAGAACATGTGATCATCCAGGCCGAGTTCTATCTGAATCTCTGACCAATCAGGC 60

Qy 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGTTCTATGTTGGATCTGGACAAAGGAG 120

Ds 61 GAGTTTATGTTGACTTTGATGTTGATGATGAGATTTCCATGTTGGATATGGCAAGAGAG 120

Qy 121 ACGGTCCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTGGAGGCTCAGGGGGCTG 180

Ds 121 ACGGTCCTGGCGCTTGAGAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAAGGTGCATTG 180

Qy 181 GCTAACATGCTATATTGAAACAAACACTTGAATACCTTGTATCCAGGCTCCACC 234

Ds 181 GCCAACATAGCTGTGGACAAAGCCACCTGGAAATCATGACAAAGCGCTCCAAC 234

```
RESULT 4
US-08-463-903-23
; Sequence 23, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: US 07/935,695
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 23
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB1/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-23

Query Match 44.9%; Score 114; DB 3; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTCTCAACTTAAGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAGACATGTGATCATCCAGCGCGAGTTCTATCTGAATCTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATCTGATCTGGACAGAGGAG 120
DB 803 GAGTTTATGTTGACTTTGATGGTATGAGATTTCCATGTGATATGCGCAAGAGGAG 862
QY 121 ACCGTCGCGATCTGGAGGAGTTTGGCCAAAGCCTTTTCCCTTGGAGGCTCAGGGCGGCTG 180
DB 863 ACGTCTGGCGGCTTGAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAGGTGCATTG 922
QY 181 GCTAACATTCGTATATTGAAACAACACTTGAATACCTTGTATGATCCAGCGTTCCACC 234
DB 923 GCCAACATAGCTGTGGACAAAGCCAACTGGAATCATGACAAAGCGCTCCAAC 976
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB1/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-23

Query Match 44.9%; Score 114; DB 3; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTCTCAACTTAAGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAGACATGTGATCATCCAGCGCGAGTTCTATCTGAATCTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATCTGATCTGGACAGAGGAG 120
DB 803 GAGTTTATGTTGACTTTGATGGTATGAGATTTCCATGTGATATGCGCAAGAGGAG 862
QY 121 ACCGTCGCGATCTGGAGGAGTTTGGCCAAAGCCTTTTCCCTTGGAGGCTCAGGGCGGCTG 180
DB 863 ACGTCTGGCGGCTTGAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAGGTGCATTG 922
QY 181 GCTAACATTCGTATATTGAAACAACACTTGAATACCTTGTATGATCCAGCGTTCCACC 234
DB 923 GCCAACATAGCTGTGGACAAAGCCAACTGGAATCATGACAAAGCGCTCCAAC 976
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB42/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-25
```

```
RESULT 5
US-08-463-903-25
; Sequence 25, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kottite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: US 07/935,695
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 25
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB48/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-27

Query Match 44.9%; Score 114; DB 3; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTCTCAACTTAAGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAGACATGTGATCATCCAGCGCGAGTTCTATCTGAATCTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATCTGATCTGGACAGAGGAG 120
DB 803 GAGTTTATGTTGACTTTGATGGTATGAGATTTCCATGTGATATGCGCAAGAGGAG 862
QY 121 ACCGTCGCGATCTGGAGGAGTTTGGCCAAAGCCTTTTCCCTTGGAGGCTCAGGGCGGCTG 180
DB 863 ACGTCTGGCGGCTTGAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAGGTGCATTG 922
QY 181 GCTAACATTCGTATATTGAAACAACACTTGAATACCTTGTATGATCCAGCGTTCCACC 234
DB 923 GCCAACATAGCTGTGGACAAAGCCAACTGGAATCATGACAAAGCGCTCCAAC 976
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB48/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-27

RESULT 6
US-08-463-903-27
; Sequence 27, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kottite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: US 07/935,695
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 27
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB48/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-27

Query Match 44.9%; Score 114; DB 3; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTCTCAACTTAAGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAGACATGTGATCATCCAGCGCGAGTTCTATCTGAATCTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATCTGATCTGGACAGAGGAG 120
DB 803 GAGTTTATGTTGACTTTGATGGTATGAGATTTCCATGTGATATGCGCAAGAGGAG 862
QY 121 ACCGTCGCGATCTGGAGGAGTTTGGCCAAAGCCTTTTCCCTTGGAGGCTCAGGGCGGCTG 180
DB 863 ACGTCTGGCGGCTTGAGAAATTTGGACGATTTGCCAGCTTTGAGGCTCAGGTGCATTG 922
QY 181 GCTAACATTCGTATATTGAAACAACACTTGAATACCTTGTATGATCCAGCGTTCCACC 234
DB 923 GCCAACATAGCTGTGGACAAAGCCAACTGGAATCATGACAAAGCGCTCCAAC 976
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRB42/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-29
```

```

; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kotite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US 08/463,903
; EARLIER APPLICATION NUMBER: US 07/935,695
; PRIOR FILING DATE: 1995-06-05
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 29
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRb41/a-FLAG construct
; LOCATION: 1..1244
US-08-463-903-29

Query Match 44.9%; Score 114; DB 3; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGCTCAACTTATGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAAGAACATGATCATCCAGCGGAGTCTATCTGAATCTTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGCAACAAGAGG 120
DB 803 GAGTTTATGTTGACTTTGATGGTGATGAGATTTCCATGTGGATATGGCAAGAGGAG 862
QY 121 ACCGTCGCGCATCTGGAGAGTTTGGCCAAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180
DB 863 ACGGTCGCGCGCTTGAAGAATTTGGACGATTTGCCAGCTTTGAGGCTCAAGGTGCAATTG 922
QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAAATCTGATCCAGCGTTCACC 234
DB 923 GCCACATAGCTGTGGACAAAGCCACCTGGAATCATGACAAAGCGTCCAAAC 976

; RESULT 9
US-07-935-695-25
; Sequence 25, Application US/07935695
; Patent No. 6329507
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kotite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/07/935,695
; CURRENT FILING DATE: 1992-08-21
; PRIOR FILING DATE: 1995-06-05
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 25
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: single chain pel B peptide-HLA-DRb42/a-FLAG construct
; LOCATION: 1..1244
; OTHER INFORMATION:
US-07-935-695-25

Query Match 44.9%; Score 114; DB 4; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGCTCAACTTATGCGCGTTTGTACAGACGCATAGACCAACAGGG 60
DB 743 ATCAAGAAGAACATGATCATCCAGCGGAGTCTATCTGAATCTTGACCAATCAGGC 802
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGCAACAAGAGG 120
DB 803 GAGTTTATGTTGACTTTGATGGTGATGAGATTTCCATGTGGATATGGCAAGAGGAG 862
QY 121 ACCGTCGCGCATCTGGAGAGTTTGGCCAAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180
DB 863 ACGGTCGCGCGCTTGAAGAATTTGGACGATTTGCCAGCTTTGAGGCTCAAGGTGCAATTG 922
QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAAATCTGATCCAGCGTTCACC 234
DB 923 GCCACATAGCTGTGGACAAAGCCACCTGGAATCATGACAAAGCGTCCAAAC 976

; RESULT 10
US-07-935-695-27
; Sequence 27, Application US/07935695
; Patent No. 6329507
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.

```

```

Query Match 44.9%; Score 114; DB 4; Length 1244;
Best Local Similarity 67.9%; Pred. No. 1.2e-29;
Matches 159; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```



US-08-462-351-4

Query Match 36.5%; Score 92.6; DB 3; Length 776;  
Best Local Similarity 67.4%; Pred. No. 2.5e-22;  
Matches 130; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 40 CAGACGCATACACCAACAGGGGAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTAT 99  
Db 126 CAGTCTCTGGAGATTTGGCCAGTACACATTTGAATTTGATGAAGATGAGATGTTCTAT 185

Qy 100 GTGATCTGGACAAAGAGAGACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCC 159  
Db 186 GTGGACTTGGATAAGAGAGACTGTCTGGATGCTTCTGAGTTTGGCCAAATTTGCGAAGC 245

Qy 160 TTTGAGGCTCAGGGCGGCTGGCTAAGATTTCTATATTGAACACAACTTGAATACCTTG 219  
Db 246 TTTGACCCCAAGGTGGACTGCAAAACATAGCTGTAGTAAACACAACTTTGGGAGCTTG 305

Qy 220 ATCCAGCGTTCCA 232  
Db 306 ACTAAGAGGTCAA 318

## RESULT 13

US-09-602-807-4  
; Sequence 4, Application US/09602807  
; Patent No. 6451314

## GENERAL INFORMATION:

; APPLICANT: Sharma, Somesh D.  
; APPLICANT: Clark, Brian R.  
; APPLICANT: Lerch, Bernard L.  
; TITLE OF INVENTION: MHC Conjugates Useful in Ameliorating  
; TITLE OF INVENTION: Autoimmunity  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA

ZIP: 94111-3834

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/602,807

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,351

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,084

FILING DATE: 30-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/590,840

FILING DATE: 23-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/969,293

FILING DATE: 14-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Parent, Annette S.

REGISTRATION NUMBER: 42,058

REFERENCE/DOCKET NUMBER: 014058-000242US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 776 base pairs

TYPE: nucleic acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: -  
; LOCATION: 1..776  
; OTHER INFORMATION: /note="I-b-A-alpha chain"

US-09-602-807-4

Query Match 36.5%; Score 92.6; DB 4; Length 776;  
Best Local Similarity 67.4%; Pred. No. 2.5e-22;  
Matches 130; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 40 CAGACGCATACACCAACAGGGGAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTAT 99  
Db 126 CAGTCTCTGGAGACATTTGGCCAGTACACATTTGAATTTGATGAAGATGAGATGTTCTAT 185

Qy 100 GTGATCTGGACAAAGAGAGACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCC 159  
Db 186 GTGGACTTGGATAAGAGAGACTGTCTGGATGCTTCTGAGTTTGGCCAAATTTGCGAAGC 245

Qy 160 TTTGAGGCTCAGGGCGGCTGGCTAAGATTTCTATATTGAACACAACTTGAATACCTTG 219  
Db 246 TTTGACCCCAAGGTGGACTGCAAAACATAGCTGTAGTAAACACAACTTTGGGAGCTTG 305

Qy 220 ATCCAGCGTTCCA 232  
Db 306 ACTAAGAGGTCAA 318

## RESULT 14

5194425-5

; Patent No. 5394425

; APPLICANT: SHARMA, SOMESH D.; LERCH, L. BERNARD; CLARK,

; BRIAN R.

; TITLE OF INVENTION: MHC-MEDIATED TOXIC CONJUGATES USEFUL IN

; AMELIORATING AUTOIMMUNITY

; NUMBER OF SEQUENCES: 9

; CURRENT APPLICATION DATA: US/07/367,751

; APPLICATION NUMBER: US/07/367,751

; FILING DATE: 21-JUN-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 210,594

; FILING DATE: 23-JUN-1988

; SEQ ID NO: 5:

; LENGTH: 776

5194425-5

Query Match 36.5%; Score 92.6; DB 6; Length 776;  
Best Local Similarity 67.4%; Pred. No. 2.5e-22;  
Matches 130; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 40 CAGACGCATACACCAACAGGGGAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTAT 99  
Db 126 CAGTCTCTGGAGACATTTGGCCAGTACACATTTGAATTTGATGAAGATGAGATGTTCTAT 185

Qy 100 GTGATCTGGACAAAGAGAGACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCC 159  
Db 186 GTGGACTTGGATAAGAGAGACTGTCTGGATGCTTCTGAGTTTGGCCAAATTTGCGAAGC 245

Qy 160 TTTGAGGCTCAGGGCGGCTGGCTAAGATTTCTATATTGAACACAACTTGAATACCTTG 219  
Db 246 TTTGACCCCAAGGTGGACTGCAAAACATAGCTGTAGTAAACACAACTTTGGGAGCTTG 305

Qy 220 ATCCAGCGTTCCA 232  
Db 306 ACTAAGAGGTCAA 318

## RESULT 15

5468481-4

; Patent No. 5468481

; APPLICANT: SHARMA, SOMESH D.; CLARK, BRIAN R.; LERCH, BERNARD L.

/ TITLE OF INVENTION: MHC CLASS II-PEPTIDE CONJUGATES USEFUL  
/ IN AMELIORATING AUTOIMMUNITY  
/ NUMBER OF SEQUENCES: 7  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/07/869,293  
/ FILING DATE: 14-APR-1992  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 690,840  
/ FILING DATE: 23-APR-1991  
/ APPLICATION NUMBER: 576,084  
/ FILING DATE: 30-AUG-1990  
/ APPLICATION NUMBER: 210,594  
/ FILING DATE: 23-JUN-1988  
/ APPLICATION NUMBER: 635,840  
/ FILING DATE: 28-DEC-1998  
/ APPLICATION NUMBER: 367,751  
/ FILING DATE: 21-JUN-1989  
/ SEQ ID NO:4:  
/ LENGTH: 776  
/ 5468481-4

Query Match 35.8%; Score 91; DB 6; Length 776;  
Best Local Similarity 66.8%; Pred. No. 9.1e-22;  
Matches 129; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 40 CAGACGCATAGACCAACAGGGGAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTAT 99  
DB 126 CAGTCTGGTGGAGACATTGGCCAGTACACATTGGAATTGATGGTGATGATGTTGTTCTAT 185  
QY 100 GTGGATCTGGACAAAGAGAGCCCTCGCATCTGGAGGAGTTTGGCCAGCCCTTTTCC 159  
DB 186 GTGGACTTGGATAAAGAGAGACTCTCTGGATGCTTCTCTGATTTTGGCCAAATTTGCCAAGC 245  
QY 160 TTTGAGGCTCAGGGGGGGTGGCTAACATTGCTATATTCAACAACTTTGAATACCTTG 219  
DB 246 TTTGACCCCAAGTGGACTGCAAAACATAGCTGTAGTAAACAACTTTGGAGTCTTG 305  
QY 220 ATCCAGCGTTCCA 232  
DB 306 ACTAAGAGGTCAA 318

RESULT 16  
US-08-596-387B-123  
; Sequence 123, Application US/08596387B  
; Patent No. 5869270  
; GENERAL INFORMATION:  
; APPLICANT: Rhode, Peter R.  
; APPLICANT: Jiao, Jin-An  
; APPLICANT: Burkhardt, Martin  
; APPLICANT: Wong, Hing  
/ TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF  
/ NUMBER OF SEQUENCES: 124  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Dade International, Inc.  
/ STREET: 1717 Deerfield Road  
/ CITY: Deerfield  
/ STATE: Illinois  
/ COUNTRY: USA  
/ ZIP: 60015  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.2, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/596,387B  
/ FILING DATE:  
/ CLASSIFICATION: 435  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/09816  
/ FILING DATE: 31-JUL-1995  
/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/382,454  
/ FILING DATE: 01-FEB-1995  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/283,302  
/ FILING DATE: 29-JUL-1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Pearson, Louise S.  
/ REGISTRATION NUMBER: 32,369  
/ REFERENCE/DOCKET NUMBER: STR-4665-CIP2  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (708) 267-5300  
/ TELEFAX: (708) 267-5376  
/ INFORMATION FOR SEQ ID NO: 123:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1382 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: unknown  
/ TOPOLOGY: unknown  
/ MOLECULE TYPE: DNA (genomic)  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 6...1382  
US-08-596-387B-123  
Query Match 34.5%; Score 87.6; DB 2; Length 1382;  
Best Local Similarity 65.2%; Pred. No. 1.9e-20;  
Matches 129; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 35 TTGTACAGACGATAGACCAACAGGGGAGTTTATGTTTGAATTGATGAAGATGAGATGT 94  
DB 853 TTTATCATGTCCTCGGAGACATTTGCCAGTACACATGAATTTGATGGTGATGATGT 912  
QY 95 TCTATGTGGATCTGGCAAGAAAGAGACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCT 154  
DB 913 TCTATGTGGATCTGGATAAGAGAAAACCTGCTGGAGCTTCTGAGTTTGGCCAAATGA 972  
QY 155 TTTCTTTGAGGCTCAGGGGGGGTGGCTAACATTGCTATATTGAACAACTTTGAATA 214  
DB 973 TACTCTTTGAGCCCAAGTGGACTGCAAAAACATAGCTGCAAGAAAACACAACTTTGGAA 1032  
QY 215 CCTTGATCCAGCGTTCCA 232  
DB 1033 TCTTGACTAAGAGGTCAA 1050

RESULT 17  
US-09-067-615-123  
; Sequence 123, Application US/09067615  
; Patent No. 6309645  
; GENERAL INFORMATION:  
; APPLICANT: Rhode, Peter R.  
; APPLICANT: Jiao, Jin-An  
; APPLICANT: Burkhardt, Martin  
; APPLICANT: Wong, Hing  
/ TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF  
/ NUMBER OF SEQUENCES: 124  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Dade International, Inc.  
/ STREET: 1717 Deerfield Road  
/ CITY: Deerfield  
/ STATE: Illinois  
/ COUNTRY: USA  
/ ZIP: 60015  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.2, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/067,615  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/596,387  
FILING DATE: 31-JUL-1995  
APPLICATION NUMBER: US 08/382,454  
FILING DATE: 01-FEB-1995  
APPLICATION NUMBER: US 08/382,454  
FILING DATE: 01-FEB-1995  
APPLICATION NUMBER: US 08/283,302  
FILING DATE: 29-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Pearson, Louise S.  
REGISTRATION NUMBER: 32,369  
REFERENCE/DOCKET NUMBER: STR-4665-CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 267-5300  
TELEFAX: (708) 267-5376  
INFORMATION FOR SEQ ID NO: 123:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1382 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6..1382

US-09-067-615-123

Query Match 34.5%; Score 87.6; DB 4; Length 1382;  
Best Local Similarity 65.2%; Pred. No. 1.9e-20;  
Matches 129; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy	35	TTGTACAGAGCGATAGACCAACAGGGAGTTTATGTTGATTTGATCAAGATGATGT	94
Db	853	TTTATCATGTCCTCGGAGACATTGGCCAGTACACACATGAATTTGATGCTGATGATGT	912
Qy	95	TCATATGAGTCTGGACAAAGAGAGACCGTCTGGCATCTGGAGAGTTTGGCCAAAGCCT	154
Db	913	TCATATGAGTCTGGATAGTAAAGAAACTGTCGGAGGCTTCCTGAGTTTGGCCAAATGA	972
Qy	155	TTTCCTTTGAGGCTCAGGGGGGCTGCTAACATGCTATATTAATTAACAAACAACTTGAATA	214
Db	973	TACTCTTTGAGCCCAAGGTGCACTGCAAAACATAGCTGCAGAAAAACACAACTTGGAA	1032
Qy	215	CCTTGATCCAGGTTCCA	232
Db	1033	TCCTGACTAAGAGTCAA	1050

RESULT 18  
PCT-US95-09816A-123  
Sequence 123, Application PC/TUS9509816A  
GENERAL INFORMATION:  
APPLICANT: Wong, King C.  
APPLICANT: Rhode, Peter R.  
APPLICANT: Widanz, Jon A.  
APPLICANT: Grammer, Susan  
APPLICANT: Edwards, Ana C.  
APPLICANT: Chavaillaz, Pierre-Andre  
APPLICANT: Jiao, Jin-An  
TITLE OF INVENTION: MEC COMPLEXES AND USES THEREOF  
NUMBER OF SEQUENCES: 123  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dade International, Inc.  
STREET: 1717 Deerfield Road  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/09816A  
FILING DATE: 31-JUL-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/382,454  
FILING DATE: 01-FEB-1995  
APPLICATION NUMBER: US 08/283,302  
FILING DATE: 29-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Pearson, Louise S.  
REGISTRATION NUMBER: 32,369  
REFERENCE/DOCKET NUMBER: STR-4665-CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 267-5300  
TELEFAX: (708) 267-5376  
INFORMATION FOR SEQ ID NO: 123:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1382 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6..1382

PCT-US95-09816A-123

Query Match 34.5%; Score 87.6; DB 5; Length 1382;  
Best Local Similarity 65.2%; Pred. No. 1.9e-20;  
Matches 129; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy	35	TTGTACAGAGCGATAGACCAACAGGGAGTTTATGTTGATTTGATCAAGATGATGT	94
Db	853	TTTATCATGTCCTCGGAGACATTGGCCAGTACACACATGAATTTGATGCTGATGATGT	912
Qy	95	TCATATGAGTCTGGACAAAGAGAGACCGTCTGGCATCTGGAGAGTTTGGCCAAAGCCT	154
Db	913	TCATATGAGTCTGGATAGTAAAGAAACTGTCGGAGGCTTCCTGAGTTTGGCCAAATGA	972
Qy	155	TTTCCTTTGAGGCTCAGGGGGGCTGCTAACATGCTATATTAATTAACAAACAACTTGAATA	214
Db	973	TACTCTTTGAGCCCAAGGTGCACTGCAAAACATAGCTGCAGAAAAACACAACTTGGAA	1032
Qy	215	CCTTGATCCAGGTTCCA	232
Db	1033	TCCTGACTAAGAGTCAA	1050

RESULT 19  
US-08-596-387B-121  
Sequence 121, Application US/08596387B  
Patent No. 5869270  
GENERAL INFORMATION:  
APPLICANT: Rhode, Peter R.  
APPLICANT: Jiao, Jin-An  
APPLICANT: Burkhardt, Martin  
APPLICANT: Wong, Hing  
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dade International, Inc.  
STREET: 1717 Deerfield Road  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,387B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09816
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pearson, Louise S.
; REGISTRATION NUMBER: 32,369
; REFERENCE/DOCKET NUMBER: STR-4665-CIP2
; TELEPHONE: (708) 267-5300
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..1382
;
US-08-596-387B-121

Query Match 34.5%; Score 87.6; DB 2; Length 1385;
Best Local Similarity 65.2%; Pred. No. 1.9e-20;
Matches 129; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 35 TTGTACAGCGCATGACCAACAGGGGAGTTTATGTTGAATTTGATGAGATGATGT 94
Db 853 TTTATCAGTCTCCTGGAGACATGGCCAGTACACACATGATTTGATGATGATGT 912
QY 95 TCTATGTGATCTGGACAGAGGAGACCGTCTGGCATCTGGAGAGTTGGCCAGCCT 154
Db 913 TCTATGTGACTTGGATAAGAGAAACTGCTGGAGGCTTCTCGAGTTGGCCAAATGA 972
QY 155 TTTCCTTTGAGGCTCAGGCGGGCTGGCTAATGCTATATTTGAACAACAACTTGAATA 214
Db 973 TACTCTTTGAGCCCCAAGGTGGACTGCAGAAACATAGCTGCAGAAAAACACAACTTGGGAA 1032
QY 215 CTTTGATCCAGCGTTCCA 232
Db 1033 TCTTGACTAAGAGGTCAA 1050

RESULT 20
PCT-US95-067-615-121
; Sequence 121, Application US/09067615
; Patent No. 6309645
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Jiao, Jin-An
; APPLICANT: Burkhardt, Martin
; APPLICANT: Wong, Hing
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: F-copy disk

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/067,615
; APPLICATION NUMBER: US/09/067,615
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/596,387
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pearson, Louise S.
; REGISTRATION NUMBER: 32,369
; REFERENCE/DOCKET NUMBER: STR-4665-CIP2
; TELEPHONE: (708) 267-5300
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..1382
;
US-09-067-615-121

Query Match 34.5%; Score 87.6; DB 4; Length 1385;
Best Local Similarity 65.2%; Pred. No. 1.9e-20;
Matches 129; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 35 TTGTACAGCGCATGACCAACAGGGGAGTTTATGTTGAATTTGATGAGATGATGT 94
Db 853 TTTATCAGTCTCCTGGAGACATTTGCCAGTACACATGATTTGATGATGATGT 912
QY 95 TCTATGTGATCTGGACAGAGGAGACCGTCTGGCATCTGGAGAGTTGGCCAGCCT 154
Db 913 TCTATGTGACTTGGATAAGAGAAACTGCTGGAGGCTTCTCGAGTTGGCCAAATGA 972
QY 155 TTTCCTTTGAGGCTCAGGCGGGCTGGCTAATGCTATATTTGAACAACAACTTGAATA 214
Db 973 TACTCTTTGAGCCCCAAGGTGGACTGCAGAAACATAGCTGCAGAAAAACACAACTTGGGAA 1032
QY 215 CTTTGATCCAGCGTTCCA 232
Db 1033 TCTTGACTAAGAGGTCAA 1050

RESULT 21
PCT-US95-09816A-121
; Sequence 121, Application PC/TUS9509816A
; GENERAL INFORMATION:
; APPLICANT: Wong, Hing C.
; APPLICANT: Rhode, Peter R.
; APPLICANT: Widanz, Jon A.
; APPLICANT: Grammer, Susan
; APPLICANT: Edwards, Ana C.
; APPLICANT: Chavaillaz, Pierre-Andre
; APPLICANT: Jiao, Jin-An
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 12:56:24 ; Search time 237 Seconds  
(without alignments)  
4795.542 Million cell updates/sec

Title: US-09-877-819b-55

Perfect score: 254

Sequence: 1 atcaaggcgaccattgtgc.....acaccaggccaccaccgat 254

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*

3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*

13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*

14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*

15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*

16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*

17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*

18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	254	100.0	254	10	US-09-877-819B-55
2	241.4	95.0	410	10	US-09-918-995-36460
3	241.4	95.0	1259	13	US-10-220-120-138
4	241.4	95.0	1348	9	US-09-925-302-351
5	241.4	95.0	1348	13	US-09-925-302-351
6	238.2	93.8	490	10	US-09-918-995-37140
7	236.6	93.1	523	15	US-10-102-524-715
8	233.8	92.0	14646	10	US-09-960-706-1043
9	233.8	92.0	14646	10	US-09-873-319-691
10	229.4	90.3	601	15	US-10-084-817-88
11	165.4	65.1	3673778	15	US-10-312-641-2
12	146	57.5	267	9	US-09-796-692-143
13	146	57.5	267	9	US-09-796-692-328
14	146	57.5	267	9	US-09-796-692-7075

15	146	57.5	267	15	US-10-040-862-143	Sequence 143, App
16	146	57.5	267	15	US-10-040-862-328	Sequence 328, App
17	146	57.5	267	15	US-10-040-862-7075	Sequence 7075, App
18	146	57.5	267	16	US-10-057-475B-143	Sequence 143, App
19	146	57.5	267	16	US-10-057-475B-328	Sequence 328, App
20	146	57.5	267	16	US-10-057-475B-7075	Sequence 7075, App
21	146	57.5	267	16	US-10-154-884B-143	Sequence 143, App
22	146	57.5	267	16	US-10-154-884B-328	Sequence 328, App
23	146	57.5	267	16	US-10-154-884B-7075	Sequence 7075, App
24	146	57.5	272	9	US-09-796-692-42	Sequence 42, Appl
25	146	57.5	272	9	US-09-796-692-3554	Sequence 3554, App
26	146	57.5	272	15	US-10-040-862-42	Sequence 42, Appl
27	146	57.5	272	15	US-10-040-862-3554	Sequence 3554, App
28	146	57.5	272	16	US-10-057-475B-42	Sequence 42, Appl
29	146	57.5	272	16	US-10-057-475B-3554	Sequence 3554, App
30	146	57.5	272	16	US-10-154-884B-42	Sequence 42, Appl
31	146	57.5	272	16	US-10-154-884B-3554	Sequence 3554, App
32	146	57.5	294	9	US-09-796-692-280	Sequence 280, App
33	146	57.5	294	9	US-09-796-692-530	Sequence 530, App
34	146	57.5	294	9	US-09-796-692-3270	Sequence 3270, App
35	146	57.5	294	15	US-10-040-862-280	Sequence 280, App
36	146	57.5	294	15	US-10-040-862-530	Sequence 530, App
37	146	57.5	294	15	US-10-040-862-3270	Sequence 3270, App
38	146	57.5	294	16	US-10-057-475B-280	Sequence 280, App
39	146	57.5	294	16	US-10-057-475B-530	Sequence 530, App
40	146	57.5	294	16	US-10-057-475B-3270	Sequence 3270, App
41	146	57.5	294	16	US-10-154-884B-280	Sequence 280, App
42	146	57.5	294	16	US-10-154-884B-530	Sequence 530, App
43	146	57.5	294	16	US-10-154-884B-3270	Sequence 3270, App
44	144.8	57.0	3673778	15	US-10-312-841-1	Sequence 1, Appl
45	115.6	45.5	641	10	US-09-847-172-43	Sequence 43, Appl
46	115.6	45.5	747	9	US-09-925-664-26	Sequence 26, Appl
47	115.6	45.5	1183	13	US-10-257-021-40	Sequence 40, Appl
48	114	44.9	369	13	US-09-878-722-75	Sequence 75, Appl
49	114	44.9	369	13	US-09-904-456-75	Sequence 75, Appl
50	114	44.9	421	10	US-09-918-995-16265	Sequence 16265, A
51	114	44.9	424	10	US-09-918-995-16429	Sequence 16429, A
52	114	44.9	444	10	US-09-918-995-11343	Sequence 11343, A
53	114	44.9	447	13	US-10-085-783A-42693	Sequence 42693, A
54	114	44.9	447	16	US-10-242-535A-42693	Sequence 42693, A
55	114	44.9	467	13	US-10-085-783A-36285	Sequence 36285, A
56	114	44.9	467	13	US-10-085-783A-42479	Sequence 42479, A
57	114	44.9	467	16	US-10-242-535A-36285	Sequence 36285, A
58	114	44.9	467	16	US-10-242-535A-42479	Sequence 42479, A
59	114	44.9	552	9	US-09-815-837-4	Sequence 4, Appl
60	114	44.9	555	9	US-09-815-837-3	Sequence 3, Appl
61	114	44.9	558	9	US-09-815-837-2	Sequence 2, Appl
62	114	44.9	561	9	US-09-815-837-5	Sequence 5, Appl
63	114	44.9	561	9	US-09-815-837-7	Sequence 7, Appl
64	114	44.9	561	9	US-09-815-837-8	Sequence 8, Appl
65	114	44.9	564	9	US-09-815-837-6	Sequence 6, Appl
66	114	44.9	570	9	US-09-815-837-9	Sequence 9, Appl
67	114	44.9	570	9	US-09-815-837-10	Sequence 10, Appl
68	114	44.9	582	9	US-09-815-837-11	Sequence 11, Appl
69	114	44.9	582	9	US-09-815-837-12	Sequence 12, Appl
70	114	44.9	600	9	US-09-815-837-13	Sequence 13, Appl
71	114	44.9	630	9	US-09-815-837-54	Sequence 54, Appl
72	114	44.9	633	9	US-09-815-837-28	Sequence 28, Appl
73	114	44.9	636	9	US-09-815-837-27	Sequence 27, Appl
74	114	44.9	639	9	US-09-815-837-26	Sequence 26, Appl
75	114	44.9	642	9	US-09-815-837-29	Sequence 29, Appl
76	114	44.9	642	9	US-09-815-837-31	Sequence 31, Appl
77	114	44.9	642	9	US-09-815-837-32	Sequence 32, Appl
78	114	44.9	642	9	US-09-815-837-33	Sequence 33, Appl
79	114	44.9	645	9	US-09-815-837-30	Sequence 30, Appl
80	114	44.9	651	9	US-09-815-837-34	Sequence 34, Appl
81	114	44.9	651	9	US-09-815-837-35	Sequence 35, Appl
82	114	44.9	663	9	US-09-815-837-36	Sequence 36, Appl
83	114	44.9	663	9	US-09-815-837-37	Sequence 37, Appl
84	114	44.9	1005	15	US-10-198-846-14059	Sequence 14059, A
85	114	44.9	1052	15	US-10-198-846-13991	Sequence 13991, A
86	114	44.9	1095	9	US-09-756-983-21	Sequence 21, Appl
87	114	44.9	1124	15	US-10-198-846-12805	Sequence 12805, A

88	114	44.9	1152	15	US-10-198-846-6766	Sequence 6766, Ap	161	84.2	33.1	1701	9	US-09-815-837-62	Sequence 62, Appl
89	114	44.9	1199	15	US-10-171-581-176	Sequence 176, Ap	162	84.2	33.1	2053	9	US-09-815-837-63	Sequence 63, Appl
90	114	44.9	1229	15	US-10-084-817-74	Sequence 74, Appl	163	84.2	33.1	2059	9	US-09-815-837-66	Sequence 66, Appl
91	114	44.9	1247	15	US-10-198-846-10318	Sequence 10318, A	164	84.2	33.1	2343	9	US-09-815-837-67	Sequence 67, Appl
92	114	44.9	1302	9	US-09-925-302-55	Sequence 55, Appl	165	84.2	33.1	2346	9	US-09-815-837-61	Sequence 61, Appl
93	114	44.9	1302	13	US-09-925-302-55	Sequence 55, Appl	166	83.4	32.8	806	15	US-10-005-626A-19	Sequence 19, Appl
94	114	44.9	1389	15	US-10-198-846-13107	Sequence 13107, A	167	83.4	32.8	927	10	US-09-971-429B-30	Sequence 30, Appl
95	113	44.5	1424	9	US-09-925-302-384	Sequence 384, App	168	83.4	32.8	956	15	US-10-037-270-229	Sequence 229, App
96	113	44.5	1424	13	US-09-925-302-384	Sequence 384, App	169	83.4	32.8	956	16	US-10-117-722-229	Sequence 229, App
97	112.4	44.3	243	14	US-10-081-281-90	Sequence 90, Appl	170	83.4	32.8	995	9	US-09-764-853-21	Sequence 21, Appl
98	112.4	44.3	621	14	US-10-081-281-90	Sequence 92, Appl	171	82.8	32.6	566	9	US-09-858-580-1	Sequence 1, Appl
99	112.4	44.3	679	15	US-10-198-846-9530	Sequence 9530, Ap	172	82.8	32.6	566	10	US-09-847-172-1	Sequence 1, Appl
100	112.4	44.3	1294	15	US-10-198-846-13553	Sequence 13553, A	173	82.2	32.4	323	9	US-09-960-352-9836	Sequence 9836, Ap
101	112.4	44.3	1463	15	US-10-198-846-13793	Sequence 12793, A	174	81.2	32.0	1075	15	US-10-198-846-6999	Sequence 6999, Ap
102	112	44.1	231	15	US-10-198-846-9513	Sequence 9513, Ap	175	80.4	31.7	441	10	US-09-918-995-14901	Sequence 14901, A
103	110.4	43.5	558	9	US-09-815-837-1	Sequence 1, Appl	176	80.4	31.7	819	15	US-10-005-626A-20	Sequence 20, Appl
104	110.4	43.5	639	9	US-09-815-837-25	Sequence 25, Appl	177	80.4	31.7	1060	10	US-09-960-706-685	Sequence 685, App
105	109.8	43.2	596	15	US-10-029-386-12122	Sequence 12122, A	178	80.4	31.7	1060	10	US-09-873-319-431	Sequence 431, App
106	109.8	43.2	1450	15	US-10-198-846-10391	Sequence 10391, A	179	80.4	31.7	1096	16	US-10-133-937-88	Sequence 88, Appl
107	109.8	43.2	5724	9	US-09-969-708-253	Sequence 253, App	180	80.4	31.7	1096	16	US-10-159-563-88	Sequence 88, Appl
108	105	41.3	644	15	US-10-198-846-6066	Sequence 6066, Ap							
109	103.8	40.9	400	9	US-09-960-352-2036	Sequence 2036, Ap							
110	102	40.2	429	15	US-10-198-846-9461	Sequence 9461, Ap							
111	101.4	39.9	449	13	US-10-085-783A-34511	Sequence 34511, A							
112	101.4	39.9	449	16	US-10-242-535A-34511	Sequence 34511, A							
113	96.4	38.0	600	9	US-09-815-837-51	Sequence 51, Appl							
114	96.4	38.0	642	9	US-09-815-837-52	Sequence 52, Appl							
115	95.6	37.6	174	15	US-10-029-386-25822	Sequence 25822, A							
116	95.4	37.6	513	9	US-09-756-692-148	Sequence 148, Ap							
117	95.4	37.6	513	9	US-09-756-692-148	Sequence 9238, Ap							
118	95.4	37.6	513	15	US-10-040-862-148	Sequence 148, App							
119	95.4	37.6	513	15	US-10-040-862-148	Sequence 9238, Ap							
120	95.4	37.6	513	16	US-10-057-475B-148	Sequence 148, App							
121	95.4	37.6	513	16	US-10-057-475B-148	Sequence 9238, Ap							
122	95.4	37.6	513	16	US-10-154-884B-148	Sequence 9238, Ap							
123	95.4	37.6	513	16	US-10-154-884B-148	Sequence 9238, Ap							
124	95.4	37.6	514	9	US-09-815-343-423	Sequence 423, App							
125	95.4	37.6	514	13	US-10-097-105-423	Sequence 423, App							
126	93.6	36.9	3673778	15	US-10-312-841-1	Sequence 1, Appl							
127	92.6	36.5	433	14	US-10-040-739-912	Sequence 912, Appl							
128	92.6	36.5	433	15	US-10-199-995-4	Sequence 4, Appl							
129	91.4	36.0	418	9	US-09-960-352-2980	Sequence 2980, Ap							
130	88	34.6	243	14	US-10-015-536-16	Sequence 16, Appl							
131	87.6	34.5	261	14	US-10-081-281-96	Sequence 96, Appl							
132	87.6	34.5	588	14	US-10-081-281-108	Sequence 108, Appl							
133	87.6	34.5	654	14	US-10-081-281-98	Sequence 98, Appl							
134	87.6	34.5	1344	14	US-10-081-281-110	Sequence 110, App							
135	87.6	34.5	1382	9	US-09-848-164-123	Sequence 123, App							
136	87.6	34.5	1382	9	US-09-900-379-123	Sequence 123, App							
137	87.6	34.5	1385	9	US-09-848-164-121	Sequence 121, App							
138	87.6	34.5	1385	9	US-09-900-379-121	Sequence 121, App							
139	87.6	34.5	1508	9	US-09-848-164-122	Sequence 122, App							
140	87.6	34.5	1508	9	US-09-766-378A-24	Sequence 24, Appl							
141	87.6	34.5	1508	9	US-09-900-379-122	Sequence 122, App							
142	86	33.9	2481	16	US-10-386-934-562	Sequence 562, App							
143	85.4	33.6	405	13	US-10-276-774-1218	Sequence 1218, Ap							
144	84.2	33.1	261	14	US-10-081-281-102	Sequence 102, App							
145	84.2	33.1	588	14	US-10-081-281-116	Sequence 116, App							
146	84.2	33.1	609	9	US-09-815-837-75	Sequence 75, Appl							
147	84.2	33.1	614	9	US-09-815-837-76	Sequence 76, Appl							
148	84.2	33.1	633	14	US-10-081-281-104	Sequence 104, App							
149	84.2	33.1	647	9	US-09-815-837-77	Sequence 77, Appl							
150	84.2	33.1	647	9	US-09-815-837-78	Sequence 78, Appl							
151	84.2	33.1	702	9	US-09-815-837-84	Sequence 84, Appl							
152	84.2	33.1	702	14	US-10-081-281-118	Sequence 118, App							
153	84.2	33.1	773	9	US-09-815-837-83	Sequence 83, Appl							
154	84.2	33.1	1013	9	US-09-815-837-85	Sequence 85, Appl							
155	84.2	33.1	1323	14	US-10-081-281-114	Sequence 114, App							
156	84.2	33.1	1662	9	US-09-815-837-80	Sequence 80, Appl							
157	84.2	33.1	1676	9	US-09-815-837-82	Sequence 82, Appl							
158	84.2	33.1	1680	9	US-09-815-837-49	Sequence 49, Appl							
159	84.2	33.1	1686	9	US-09-815-837-65	Sequence 65, Appl							
160	84.2	33.1	1698	9	US-09-815-837-79	Sequence 79, Appl							

## ALIGNMENTS

## RESULT 1

US-09-877-819B-55  
; Sequence 55, Application US/09877819B  
; Publication No. US20030190609A1  
; GENERAL INFORMATION:  
; APPLICANT: Torney, David  
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing  
; FILE REFERENCE: S-94,864  
; CURRENT APPLICATION NUMBER: US/09/877,819B  
; CURRENT FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 55  
; LENGTH: 254  
; TYPE: DNA  
; ORGANISM: Human HLA  
US-09-877-819B-55

Query Match 100.0%; Score 254; DB 10; Length 254;  
Best Local Similarity 100.0%; Pred. No. 2.5e-77; Indels 0; Gaps 0;  
Matches 254; Conservative 0; Mismatches 0;

QY 1 ATCAAGCGGACCATGTGTCAACTTATCCGCGTTTGTACAGCCATAGACCAACAGGG 60  
1 ATCAAGCGGACCATGTGTCAACTTATCCGCGTTTGTACAGCCATAGACCAACAGGG 60

QY 51 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTCGAAGAAGAG 120  
51 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTCGAAGAAGAG 120

QY 121 ACCGCTGGCATCTGGAGAGTTTGGCAAGCTTTTCCTTGGCTCAGGCGGGCTG 180  
121 ACCGCTGGCATCTGGAGAGTTTGGCAAGCTTTTCCTTGGCTCAGGCGGGCTG 180

QY 181 GCTAACATTGTTATTTGAACAACTTGATACCTTGATCCAGCGTTCCACCACTC 240  
181 GCTAACATTGTTATTTGAACAACTTGATACCTTGATCCAGCGTTCCACCACTC 240

QY 241 AGGCCACCAACCGAT 254  
241 AGGCCACCAACCGAT 254

## RESULT 2

US-09-918-995-36460  
; Sequence 36460, Application US/09918995  
; Publication No. US2003037623A1

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36460
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-36460

Query Match      95.0%; Score 241.4; DB 10; Length 410;
Best Local Similarity 99.2%; Pred. No. 7.4e-73;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy      1 ATCAAGCGGACCAATGTCACACTTATGCCCGTTTGTACAGACGCATAGACCAACAGGG 60
Db      145 ATCAAGCGGACCAATGTCACACTTATGCCCGTTTGTACAGACGCATAGACCAACAGGG 204
Qy      61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTCGACAAAGAGGAG 120
Db      205 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTCGACAAAGAGGAG 264
Qy      121 ACCGTCTGGCATCTGGAGGAGTTTGGCCAGCCCTTTCTTTGAGGCTCAGGCGGGGCTG 180
Db      265 ACCGTCTGGCATCTGGAGGAGTTTGGCCAGCCCTTTCTTTGAGGCTCAGGCGGGGCTG 324
Qy      181 GCTAACATTGCTATATTGAACAACAACCTTGTAATACCTTGATCCAGCGTTCC-ACCACT 239
Db      325 GCTAACATTGCTATATTGAACAACAACCTTGTAATACCTTGATCCAGCGTTCCACCACT 384
Qy      240 CAGGCCACCAACCGAT 254
Db      385 CAGGCCACCAACCGAT 399

RESULT 3
US-10-220-120-138
; Sequence 138 Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CEALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce E.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFHO, Abel
; APPLICANT: WRIGHT, Rachel J.

; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,813; 60/184,773; 60/184,776;
60/184,693; 60/184,771; 60/184,837; 60/184,697; 60/184,841;
60/184,769; 60/184,768; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 138
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:1169865.1:2000MAY01
US-10-220-120-138

Query Match      95.0%; Score 241.4; DB 13; Length 1259;
Best Local Similarity 99.2%; Pred. No. 1.4e-72;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy      1 ATCAAGCGGACCAATGTCACACTTATGCCCGTTTGTACAGACGCATAGACCAACAGGG 60
Db      308 ATCAAGCGGACCAATGTCACACTTATGCCCGTTTGTACAGACGCATAGACCAACAGGG 367
Qy      61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGACAAAGAGGAG 120
Db      368 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGACAAAGAGGAG 427
Qy      121 ACCGTCTGGCATCTGGAGGAGTTTGGCCAGCCCTTTCTTTGAGGCTCAGGCGGGGCTG 180
Db      428 ACCGTCTGGCATCTGGAGGAGTTTGGCCAGCCCTTTCTTTGAGGCTCAGGCGGGGCTG 487
Qy      181 GCTAACATTGCTATATTGAACAACAACCTTGTAATACCTTGATCCAGCGTTCC-ACCACT 239
Db      488 GCTAACATTGCTATATTGAACAACAACCTTGTAATACCTTGATCCAGCGTTCCACCACT 547
Qy      240 CAGGCCACCAACCGAT 254
Db      548 CAGGCCACCAACCGAT 562

RESULT 4
US-09-925-302-351
; Sequence 351, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
```

;; PRIOR APPLICATION NUMBER: PCT/US00/05918  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 896  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 351  
;; LENGTH: 1348  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (12)  
;; NAME/KEY: misc feature  
;; LOCATION: (12)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc feature  
;; LOCATION: (24)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc feature  
;; LOCATION: (36)  
;; NAME/KEY: misc feature  
;; LOCATION: (1294)  
;; NAME/KEY: misc feature  
;; LOCATION: (1307)  
;; NAME/KEY: misc feature  
;; LOCATION: (1318)  
;; NAME/KEY: misc feature  
;; LOCATION: (1318)  
;; NAME/KEY: misc feature  
;; LOCATION: (1329)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-302-351

Query Match 95.0%; Score 241.4; DB 9; Length 1348;  
Best Local Similarity 99.2%; Pred. No. 1.4e-72;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCACCTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 60  
Db 246 ATCAAGCGGACCATGTGTCACCTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 305  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 120  
Db 306 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 365  
QY 121 ACCGCTCGCATCTGAGAGTTTGGCCAGCGTTTCTTGTGAGGCTCAGGGCGGCTG 180  
Db 366 ACCGCTCGCATCTGAGAGTTTGGCCAGCGTTTCTTGTGAGGCTCAGGGCGGCTG 425  
QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
Db 426 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCCACACACT 485  
QY 240 CAGGCCACCAACCGAT 254  
Db 486 CAGGCCACCAACCGAT 500

## RESULT 5

US-09-925-302-351  
;; Sequence 351, Application US/09925302  
;; Publication No. US20030064072A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
;; FILE REFERENCE: PA104  
;; CURRENT APPLICATION NUMBER: US/09/925,302  
;; CURRENT FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05918  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12

;; NUMBER OF SEQ ID NOS: 896  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 351  
;; LENGTH: 1348  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (12)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc feature  
;; LOCATION: (124)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc feature  
;; LOCATION: (36)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc feature  
;; LOCATION: (1294)  
;; NAME/KEY: misc feature  
;; LOCATION: (1307)  
;; NAME/KEY: misc feature  
;; LOCATION: (1318)  
;; NAME/KEY: misc feature  
;; LOCATION: (1329)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-302-351

Query Match 95.0%; Score 241.4; DB 13; Length 1348;  
Best Local Similarity 99.2%; Pred. No. 1.4e-72;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCACCTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 60  
Db 246 ATCAAGCGGACCATGTGTCACCTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 305  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 120  
Db 306 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 365  
QY 121 ACCGCTCGCATCTGAGAGTTTGGCCAGCGTTTCTTGTGAGGCTCAGGGCGGCTG 180  
Db 366 ACCGCTCGCATCTGAGAGTTTGGCCAGCGTTTCTTGTGAGGCTCAGGGCGGCTG 425  
QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
Db 426 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCCACACACT 485  
QY 240 CAGGCCACCAACCGAT 254  
Db 486 CAGGCCACCAACCGAT 500

## RESULT 6

US-09-918-995-37140  
;; Sequence 37140, Application US/09918995  
;; Publication No. US20030073623A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Hyseq, Inc.  
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
;; FILE REFERENCE: 20411-756  
;; CURRENT APPLICATION NUMBER: US/09/918,995  
;; CURRENT FILING DATE: 2001-07-30  
;; PRIOR APPLICATION NUMBER: US/09/235,076  
;; PRIOR FILING DATE: 1999-01-20  
;; NUMBER OF SEQ ID NOS: 38054  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 37140  
;; LENGTH: 490  
;; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-919-995-37140

Query Match      93.8%; Score 238.2; DB 10; Length 490;
Best Local Similarity 98.4%; Pred. No. 1e-71;
Matches 251; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATCAAGGGGACCACTGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGG 60
DB 218 ATCAAGGGGACCACTGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGG 277
QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
DB 278 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 337
QY 121 ACCGCTCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180
DB 338 ACCGCTCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 397
QY 181 GCTACATTCCTATATTGAACAACACTTCAATACCTTATCCAGGTTCC-ACCACACT 239
DB 398 GCTACATTCCTATATTGAACAACACTTCAATACCTTATCCAGGTTCC-ACCACACT 457
QY 240 CAGGCCACCAACCGAT 254
DB 458 CAGGCCACCAACCGAT 472

RESULT 7
US-10-102-524-715
; Sequence 715, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 715
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-715

Query Match      93.1%; Score 236.6; DB 15; Length 523;
Best Local Similarity 98.0%; Pred. No. 3.9e-71;
Matches 250; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATCAAGGGGACCACTGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGG 60
DB 133 ATCAAGGGGACCACTGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGG 192
QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
DB 193 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 252
QY 121 ACCGCTCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 180
DB 253 ACCGCTCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTG 312
QY 181 GCTACATTCCTATATTGAACAACACTTCAATACCTTATCCAGGTTCC-ACCACACT 239

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-960-706-1043
; Sequence 1043, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)-(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-960-706-1043

Query Match      92.0%; Score 233.8; DB 10; Length 14646;
Best Local Similarity 98.8%; Pred. No. 2.2e-69;
Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 7 GCGGACCATGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGGAGTTT 66
DB 6333 GCGGACCATGTCTCACTTATCCCGGTTGTACAGACCATAGACCAACAGGGAGTTT 6392
QY 67 ATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAGACCGTC 126
DB 6393 ATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAGACCGTC 6452
QY 127 TGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTGGCTAAC 186
DB 6453 TGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCCTTTGAGGCTCAGGCGGGCTGGCTAAC 6512
QY 187 ATTGCTATATTGAACAACACTTGAATACCTTGAATCCAGGTTCC-ACCACACTCAGGCC 245
DB 6513 ATTGCTATATTGAACAACACTTGAATACCTTGAATCCAGGTTCC-ACCACACTCAGGCC 6572
QY 246 ACCACCGAT 254
DB 6573 ACCACCGAT 6581

RESULT 9
US-09-873-319-691
; Sequence 691, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US

```

; CURRENT APPLICATION NUMBER: US/09/873,319A  
 ; CURRENT FILING DATE: 2001-06-05  
 ; EARLIER APPLICATION NUMBER: US 60/223,323  
 ; EARLIER FILING DATE: 2000-08-07  
 ; NUMBER OF SEQ ID NOS: 755  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 691  
 ; LENGTH: 14646  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X03100  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(14646)  
 ; OTHER INFORMATION: n = a or c or g or t  
 US-09-873-319-691

Query Match 92.0%; Score 233.8; DB 10; Length 14646;  
 Best Local Similarity 98.8%; Pred. No. 2.2e-69;  
 Matches 246; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 7 GCGGACCATGTGTCACCTTATCCCGGTTGTACAGACCATAGACCAACAGGGAGTTT 66  
 Db 6333 GCGGACCATGTGTCACCTTATCCCGGTTGTACAGACCATAGACCAACAGGGAGTTT 6392  
 QY 67 ATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAGACCGTC 126  
 Db 6393 ATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAGACCGTC 6452  
 QY 127 TGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGGCGGCTGGCTAAC 186  
 Db 6453 TGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGGCGGCTGGCTAAC 6512  
 QY 187 ATGCTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTCC-ACCACACTCAGGCC 245  
 Db 6513 ATGCTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTCC-ACCACACTCAGGCC 6572  
 QY 246 ACCACCGAT 254  
 Db 6573 ACCACCGT 6581

RESULT 10  
 US-10-084-817-88  
 ; Sequence 88, Application US/10084817  
 ; Publication No. US20030119039A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Susan Stuart  
 ; APPLICANT: Jed G. Nuchtern  
 ; APPLICANT: Sharon E. Plon  
 ; APPLICANT: Jason M. Shohet  
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
 ; CURRENT APPLICATION NUMBER: US/10/084,817  
 ; CURRENT FILING DATE: 2002-02-25  
 ; PRIOR APPLICATION NUMBER: 60/270,784  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 365  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 88  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 1090035.1  
 US-10-084-817-88

Query Match 90.3%; Score 229.4; DB 15; Length 601;  
 Best Local Similarity 98.8%; Pred. No. 1.3e-68;  
 Matches 252; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 ATCAAGGCGGACCATGTGTCACCTTATCCCGGTTGTACAGACCATAGACCAACAGGG 60  
 Db 166 ATCAAGGCGGACCATGTGTCACCTTATCCCGGTTGTACAGACCATAGACCAACAGGG 225  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAG 120  
 Db 226 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAG 285  
 QY 121 ACCGTCGTCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGGCGGCTG 180  
 Db 286 ACCGTCGTCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGGCGGCTG 344  
 QY 181 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTCC-ACCACACT 239  
 Db 345 GCTAACATTGCTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGTCC-ACCACACT 404  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 405 CAGGCCACCAACCGAT 419

RESULT 11  
 US-10-312-841-2  
 ; Sequence 2, Application US/10312841  
 ; Publication No. US20030186277A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Epigenomics AG  
 ; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
 ; FILE REFERENCE: E01/1208/WO  
 ; CURRENT APPLICATION NUMBER: US/10/312,841  
 ; CURRENT FILING DATE: 2002-12-30  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SEQ ID NO 2  
 ; LENGTH: 3673778  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; FEATURE: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; NAME/KEY: unsure  
 ; LOCATION: (379615)  
 US-10-312-841-2

Query Match 65.1%; Score 165.4; DB 15; Length 3673778;  
 Best Local Similarity 83.9%; Pred. No. 2e-44;  
 Matches 187; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 7 GCGGACCATGTGTCACCTTATCCCGGTTGTACAGACCATAGACCAACAGGGAGTTT 66  
 Db 375717 GCGGATTATGTGTTAAATTTATGTCGCGTTGCTATAGACGTATAGATTAAATAGGGAGTTT 375776  
 QY 67 ATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAGACCGTC 126  
 Db 375777 ATGTTTGAATTTGATGAAGATGAGATGTTTCTATGTGGATCTCGACAAGAGGAGATCGTT 375836  
 QY 127 TGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGGCGGCTGGCTAAC 186  
 Db 375837 TGGTATTGGAGGAGTTTGGTAAAGTTTCTTTTGGGTTTACGGCGGTTGGTTAAT 375896  
 QY 187 ATTGCTATATTGAACAACAACCTTGAATACCTTGTATCCAGCGT 229  
 Db 375897 ATTGCTATATTGAATAATAATTGAATATTGTTAGCGTT 375939

RESULT 12  
 US-09-796-692-143  
 ; Sequence 143, Application US/09796692  
 ; Publication No. US20020198362A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: Mannion, Jane  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY

1 TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
2 FILE REFERENCE: 2077.001200  
3 CURRENT APPLICATION NUMBER: US/09/796,692  
4 CURRENT FILING DATE: 2001-03-01  
5 PRIOR APPLICATION NUMBER: 60/186,126  
6 PRIOR FILING DATE: 2000-03-01  
7 PRIOR APPLICATION NUMBER: 60/190,479  
8 PRIOR FILING DATE: 2000-03-17  
9 PRIOR APPLICATION NUMBER: 60/200,545  
10 PRIOR FILING DATE: 2000-04-27  
11 PRIOR APPLICATION NUMBER: 60/200,303  
12 PRIOR FILING DATE: 2000-04-28  
13 PRIOR APPLICATION NUMBER: 60/200,779  
14 PRIOR FILING DATE: 2000-04-28  
15 PRIOR APPLICATION NUMBER: 60/200,999  
16 PRIOR FILING DATE: 2000-05-01  
17 PRIOR APPLICATION NUMBER: 60/202,084  
18 PRIOR FILING DATE: 2000-05-04  
19 PRIOR APPLICATION NUMBER: 60/206,201  
20 PRIOR FILING DATE: 2000-05-01  
21 PRIOR APPLICATION NUMBER: 60/200,303  
22 PRIOR FILING DATE: 2000-07-14  
23 PRIOR APPLICATION NUMBER: 60/222,903  
24 PRIOR FILING DATE: 2000-08-03  
25 PRIOR APPLICATION NUMBER: 60/223,416  
26 PRIOR FILING DATE: 2000-08-04  
27 PRIOR APPLICATION NUMBER: 60/223,378  
28 PRIOR FILING DATE: 2000-08-07  
29 NUMBER OF SEQ ID NOS: 9597  
30 SOFTWARE: FastSeq for Windows Version 3.0  
31 SEQ ID NO 143  
32 LENGTH: 267  
33 TYPE: DNA  
34 ORGANISM: Homo sapiens  
35 US-09-796-692-143

Query Match 57.5%; Score 146; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 6.2e-40;  
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATCAAGCGGACCAATGTGTCAACTTATGCCCGCTTTGTACAGACGCATAGACCAACAGGG 60  
Db 122 ATCAAGCGGACCAATGTGTCAACTTATGCCCGCTTTGTACAGACGCATAGACCAACAGGG 181  
Qy 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGTTCTATGTCGATCTGGACAGAGAG 120  
Db 182 GAGTTTATGTTGAAATTTGATGAGATGAGATGTTCTATGTCGATCTGGACAGAGAG 241  
Qy 121 ACCGTCGTCATCTGGAGGAGTTTGG 146  
Db 242 ACCGTCGTCATCTGGAGGAGTTTGG 267

RESULT 13  
US-09-796-692-328  
Sequence 328, Application US/09/796,692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303

1 PRIOR FILING DATE: 2000-04-28  
2 PRIOR APPLICATION NUMBER: 60/200,779  
3 PRIOR FILING DATE: 2000-04-28  
4 PRIOR APPLICATION NUMBER: 60/200,999  
5 PRIOR FILING DATE: 2000-05-01  
6 PRIOR APPLICATION NUMBER: 60/202,084  
7 PRIOR FILING DATE: 2000-05-04  
8 PRIOR APPLICATION NUMBER: 60/206,201  
9 PRIOR FILING DATE: 2000-05-22  
10 PRIOR APPLICATION NUMBER: 60/218,950  
11 PRIOR FILING DATE: 2000-07-14  
12 PRIOR APPLICATION NUMBER: 60/222,903  
13 PRIOR FILING DATE: 2000-08-03  
14 PRIOR APPLICATION NUMBER: 60/223,416  
15 PRIOR FILING DATE: 2000-08-04  
16 PRIOR APPLICATION NUMBER: 60/223,378  
17 PRIOR FILING DATE: 2000-08-07  
18 NUMBER OF SEQ ID NOS: 9597  
19 SOFTWARE: FastSeq for Windows Version 3.0  
20 SEQ ID NO 328  
21 LENGTH: 267  
22 TYPE: DNA  
23 ORGANISM: Homo sapiens  
24 US-09-796-692-328  
Query Match 57.5%; Score 146; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 6.2e-40;  
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATCAAGCGGACCAATGTGTCAACTTATGCCCGCTTTGTACAGACGCATAGACCAACAGGG 60  
Db 122 ATCAAGCGGACCAATGTGTCAACTTATGCCCGCTTTGTACAGACGCATAGACCAACAGGG 181  
Qy 61 GAGTTTATGTTGAAATTTGATGAGATGAGATGTTCTATGTCGATCTGGACAGAGAG 120  
Db 182 GAGTTTATGTTGAAATTTGATGAGATGAGATGTTCTATGTCGATCTGGACAGAGAG 241  
Qy 121 ACCGTCGTCATCTGGAGGAGTTTGG 146  
Db 242 ACCGTCGTCATCTGGAGGAGTTTGG 267

RESULT 14  
US-09-796-692-7075  
Sequence 7075, Application US/09/796,692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Paul A.  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14

US-10-040-862-328



```
Query Match          57.5%; Score 146; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 6.2e-40;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 60
DB 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 181

QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120
DB 182 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 241

QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGG 146
DB 242 ACCGTCCTGGCATCTGGAGGAGTTTGG 267

RESULT 17
US-10-040-862-7075
; Sequence 7075, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-0135200S
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7075

Query Match          57.5%; Score 146; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 6.2e-40;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 60
DB 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 181
```

```
QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120
DB 182 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 241

QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGG 146
DB 242 ACCGTCCTGGCATCTGGAGGAGTTTGG 267

RESULT 18
US-10-057-475B-143
; Sequence 143, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-0144020S
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-143

Query Match          57.5%; Score 146; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 6.2e-40;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 60
DB 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTGTACAGACGCATAGACCAACAGGG 181

QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120
DB 182 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 241

QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGG 146
DB 242 ACCGTCCTGGCATCTGGAGGAGTTTGG 267
```

## RESULT 19

US-10-057-475B-328  
; Sequence 328, Application US/10057475B  
; Publication No. US20040002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Clapper, Jonathan David  
; APPLICANT: Wang, Aijun  
; APPLICANT: Ordonez, Nadia  
; APPLICANT: Carter, Lauren  
; APPLICANT: McNeill, Patricia Dianne  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-014402US  
; CURRENT APPLICATION NUMBER: US/10/057,475B  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 328  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-057-475B-328

Query Match 57.5%; Score 146; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 6.2e-40;  
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
Db 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 181  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 120  
Db 182 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 241  
QY 121 ACCGTCTGGCATCTGGAGGAGTTGG 146  
Db 242 ACCGTCTGGCATCTGGAGGAGTTGG 267  
RESULT 20  
US-10-057-475B-7075  
; Sequence 7075, Application US/10057475B  
; Publication No. US20040002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander

Query Match 57.5%; Score 146; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 6.2e-40;  
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
Db 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 181  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 120  
Db 182 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 241  
QY 121 ACCGTCTGGCATCTGGAGGAGTTGG 146  
Db 242 ACCGTCTGGCATCTGGAGGAGTTGG 267

## RESULT 20

US-10-057-475B-7075  
; Sequence 7075, Application US/10057475B  
; Publication No. US20040002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Clapper, Jonathan David  
; APPLICANT: Wang, Aijun  
; APPLICANT: Ordonez, Nadia  
; APPLICANT: Carter, Lauren  
; APPLICANT: McNeill, Patricia Dianne  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-014402US  
; CURRENT APPLICATION NUMBER: US/10/057,475B  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7075  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-057-475B-7075  
Query Match 57.5%; Score 146; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 6.2e-40;  
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
Db 122 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 181  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 120  
Db 182 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAGAGGAG 241  
QY 121 ACCGTCTGGCATCTGGAGGAGTTGG 146  
Db 242 ACCGTCTGGCATCTGGAGGAGTTGG 267  
RESULT 21  
US-10-154-884B-143  
; Sequence 143, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 12:34:14 ; Search time 1537 Seconds  
(without alignments)

4934.934 Million cell updates/sec

Title: US-09-877-819B-55

Perfect score: 254

Sequence: 1 atcaagcgaccatgtgc.....acactcagccaccaccgat 254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rtd:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gsl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.4	95.0	437	12	BM767805
2	241.4	95.0	484	10	BF819626
3	241.4	95.0	491	12	BM694247
4	241.4	95.0	514	14	CD698819
					CD698819 EST15342

5	241.4	95.0	515	12	BM769742
6	241.4	95.0	519	12	BM541135
7	241.4	95.0	540	14	CD695435
8	241.4	95.0	545	13	BM783326
9	241.4	95.0	557	12	BM535978
10	241.4	95.0	563	14	BM266399
11	241.4	95.0	571	14	CA942442
12	241.4	95.0	579	12	BM831052
13	241.4	95.0	582	12	BM737984
14	241.4	95.0	585	9	AV706521
15	241.4	95.0	585	13	BM783392
16	241.4	95.0	588	12	BM783392
17	241.4	95.0	592	12	BM783392
18	241.4	95.0	594	12	BM783392
19	241.4	95.0	595	14	CD693703
20	241.4	95.0	597	14	CD693703
21	241.4	95.0	599	13	BM783392
22	241.4	95.0	601	10	BM783392
23	241.4	95.0	602	13	BM783392
24	241.4	95.0	610	14	CA405960
25	241.4	95.0	626	12	BM783392
26	241.4	95.0	627	14	CD693703
27	241.4	95.0	635	12	BM783392
28	241.4	95.0	643	12	BM783392
29	241.4	95.0	645	12	BM783392
30	241.4	95.0	645	12	BM783392
31	241.4	95.0	659	14	CD687077
32	241.4	95.0	667	9	AL048565
33	241.4	95.0	669	13	BM783392
34	241.4	95.0	676	12	BM783392
35	241.4	95.0	685	10	BM783392
36	241.4	95.0	724	10	BM783392
37	241.4	95.0	724	10	BM783392
38	241.4	95.0	734	9	AV733676
39	241.4	95.0	743	14	CD104670
40	241.4	95.0	743	14	CD104670
41	241.4	95.0	745	14	CD984378
42	241.4	95.0	758	14	CD984378
43	241.4	95.0	789	12	BM783392
44	241.4	95.0	791	13	BM783392
45	241.4	95.0	802	12	BM783392
46	241.4	95.0	806	14	BM783392
47	241.4	95.0	811	12	BM783392
48	241.4	95.0	821	12	BM783392
49	241.4	95.0	826	12	BM783392
50	241.4	95.0	830	14	BM783392
51	241.4	95.0	832	14	BM783392
52	241.4	95.0	837	12	BM783392
53	241.4	95.0	843	12	BM783392
54	241.4	95.0	851	14	CD558780
55	241.4	95.0	857	13	BM783392
56	241.4	95.0	871	12	BM783392
57	241.4	95.0	872	13	BM783392
58	241.4	95.0	873	12	BM783392
59	241.4	95.0	878	14	BM783392
60	241.4	95.0	880	9	AV139061
61	241.4	95.0	881	14	CD558780
62	241.4	95.0	889	14	CD558780
63	241.4	95.0	901	13	BM783392
64	241.4	95.0	915	10	BM783392
65	241.4	95.0	917	12	BM783392
66	241.4	95.0	923	13	BM783392
67	241.4	95.0	926	13	BM783392
68	241.4	95.0	943	13	BM783392
69	241.4	95.0	959	10	BM783392
70	241.4	95.0	960	13	BM783392
71	241.4	95.0	967	12	BM783392
72	241.4	95.0	980	10	BM783392
73	241.4	95.0	980	13	BM783392
74	241.4	95.0	990	12	BM783392
75	241.4	95.0	1005	13	BM783392
76	241.4	95.0	1007	13	BM783392
77	241.4	95.0	1015	13	BM783392

78	241.4	95.0	1025	12	BQ054502	AGENCOURT	151	227	89.4	864	10	BF663966	602145632
79	241.4	95.0	1026	13	BQ061344	AGENCOURT	152	227	89.4	874	12	BQ049475	AGENCOURT
80	241.4	95.0	1027	13	BQ064032	AGENCOURT	153	227	89.4	895	10	BF794146	602255521
81	241.4	95.0	1030	13	BQ063085	AGENCOURT	154	227	89.4	906	12	BF517914	603041396
82	241.4	95.0	1061	12	BQ054682	AGENCOURT	155	227	89.4	910	10	BF797420	602257071
83	241.4	95.0	1108	13	BQ072806	AGENCOURT	156	227	89.4	912	12	BF765449	603050505
84	241.4	95.0	1111	14	CD517173	AGENCOURT	157	227	89.4	916	10	BF794242	602255439
85	241.4	95.0	1201	13	BK363746	AGENCOURT	158	227	89.4	919	9	AL566694	AL566694
86	241.4	95.0	1217	12	RM544066	AGENCOURT	159	227	89.4	922	10	BF794550	602255814
87	239.8	94.4	469	12	BM766631	K-EST0048	160	227	89.4	957	10	BF796372	602260007
88	239.8	94.4	526	14	CD705087	AGENCOURT	161	227	89.4	962	10	BF797647	602257466
89	239.8	94.4	545	13	BQ267707	AGENCOURT	162	227	89.4	989	12	BQ049982	AGENCOURT
90	239.8	94.4	629	14	CD691540	AGENCOURT	163	227	89.4	991	12	BM904334	AGENCOURT
91	239.8	94.4	658	9	AV763677	AGENCOURT	164	227	89.4	1032	10	BQ050343	AGENCOURT
92	239.8	94.4	672	14	CD692959	AGENCOURT	165	227	89.4	1043	13	BQ882939	AGENCOURT
93	239.8	94.4	684	14	CD693272	AGENCOURT	166	227	89.4	1094	12	BF756707	603024543
94	239.8	94.4	709	12	BG758164	602712334	167	227	89.4	1106	12	BM477269	AGENCOURT
95	239.8	94.4	743	14	CB956955	AGENCOURT	168	227	89.4	1148	12	BM456056	AGENCOURT
96	239.8	94.4	809	14	CB992720	AGENCOURT	169	227	89.4	1201	13	EX397131	EX397131
97	239.8	94.4	860	12	BG755507	602713982	170	227	89.4	1361	12	BM920808	AGENCOURT
98	239.8	94.4	964	13	BQ642742	AGENCOURT	171	226.8	89.3	415	9	AA838010	AA838010
99	238.2	93.8	547	10	AW406086	UI-HF-BL0	172	226.6	89.2	959	13	BQ057748	BQ057748
100	238.2	93.8	559	14	CD700948	EST17504	173	226.6	89.2	1081	13	EX399078	EX399078
101	238.2	93.8	706	12	BG431973	602498115	174	225.8	88.9	557	14	CD687141	CD687141
102	236.8	93.2	854	12	BG756171	603713472	175	225.4	88.7	408	12	BM834672	BM834672
103	235.2	92.6	790	12	B1837330	603090674	176	225.4	88.7	444	14	CD704753	CD704753
104	235	92.5	671	12	BG540027	602568650	177	225.4	88.7	527	14	CD684349	CD684349
105	235	92.5	708	14	N250399	YX18d12.r1	178	225.4	88.7	598	14	CD691317	CD691317
106	235	92.5	766	12	BG541205	602569986	179	225.4	88.7	872	12	BG542978	BG542978
107	233.8	92.0	732	12	BG699921	602681285	180	225.2	88.7	443	14	CD687150	CD687150
108	233.4	91.9	618	14	CD686880	EST3401.h							
109	232.2	91.4	387	14	CB267061	1005967.H							
110	230.2	90.6	732	14	CB997879	AGENCOURT							
111	230.2	90.6	823	14	CB931431	AGENCOURT							
112	229.4	90.3	702	12	BG541647	602571219							
113	229.4	90.3	720	10	BF976100	602244336							
114	229.4	90.3	776	12	B1227278	602948666							
115	229.4	90.3	859	12	BG546165	602573582							
116	229.4	90.3	884	12	B1545349	603187481							
117	229.4	90.3	890	12	BG754449	602710081							
118	228.8	90.1	544	10	B8740055	601484433							
119	228.8	90.1	547	14	CD706950	EST23477							
120	228.8	90.1	786	12	BG427527	602494551							
121	228.6	90.0	494	14	CD698716	EST15239							
122	228.6	90.0	533	14	CD687511	EST4032.h							
123	228.6	90.0	534	12	BG756165	602713466							
124	228.6	90.0	546	10	AW353777	AGENCOURT							
125	228.6	90.0	602	14	CD699993	EST16517							
126	228.6	90.0	663	14	CD701461	EST17985							
127	228.6	90.0	693	12	BG685335	602637767							
128	228.6	90.0	796	14	CB997069	AGENCOURT							
129	228.6	90.0	797	12	BG758487	602712686							
130	228.6	90.0	817	12	BG758419	602712617							
131	228.6	90.0	843	10	BF75896	602246288							
132	228.6	90.0	879	13	BUS29093	AGENCOURT							
133	228.6	90.0	930	13	BQ712102	AGENCOURT							
134	228.6	90.0	966	13	BQ706261	AGENCOURT							
135	228.6	90.0	998	13	AW899275	AGENCOURT							
136	227	89.4	439	10	AW406315	UI-HF-BL0							
137	227	89.4	483	14	CD692032	EST1555.h							
138	227	89.4	530	14	CD708787	EST25314							
139	227	89.4	548	14	CD696409	EST12932							
140	227	89.4	588	14	CD698528	EST15051							
141	227	89.4	590	14	CD707303	EST23830							
142	227	89.4	605	12	BM820452	K-EST0088							
143	227	89.4	608	14	CD706661	EST23188							
144	227	89.4	628	14	CD702973	EST19580							
145	227	89.4	646	14	CD689742	EST6265.h							
146	227	89.4	698	12	B1834690	603090472							
147	227	89.4	700	14	CD696900	EST13423							
148	227	89.4	799	9	AV752267	AV752267							
149	227	89.4	851	12	B1769021	603058180							
150	227	89.4	858	12	B1833342	603087948							

## ALIGNMENTS

RESULT 1	LOCUS	BM767805	BM767805	437 bp	mrna	linear	EST 04-MAR-2002
BM767805	DEFINITION	K-EST0050302	S1SNUS2	Homo sapiens	cdna	clone	S1SNUS2-12-E02 5',
	ACCESSION	BM767805			mRNA sequence.		
	VERSION	BM767805					
	KEYWORDS	EST.					
	SOURCE	BM767805.1	GI:19097420				
	ORGANISM	Homo sapiens	(human)				
	REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	AUTHORS	Mammalia; Rutheria; Primates; Catarrhini; Homiidae; Homo.					
		Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,					
		Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and					
		Kim, Y.S.					
	TITLE	21C Frontier Korean EST Project 2001					
	JOURNAL	Unpublished (2002)					
	COMMENT	Contact: Kim YS					
		Genome Research Center					
		Korea Research Insitute of Bioscience & Biotechnology					
		52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea					
		Tel: +82-42-860-4470					
		Fax: +82-42-860-4409					
		Email: yongsung@mail.kribb.re.kr					
		Plate: 12 row: E column: 02					
		High quality sequence stop: 437.					
	FEATURES	Location/Qualifiers					
	source	1. 437					
		/organism="Homo sapiens"					
		/mol_type="mRNA"					
		/db_xref="taxon:9606"					
		/clone="S1SNUS2-12-E02"					
		/sex="F"					
		/tissue_type="Ascites"					
		/cell_type="Lymphoblast-like"					
		/cell_line="SNU-5"					
		/lab_host="Top10", "					

/clone lib="S15NUS2"  
 /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F' by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library. After analyzing and  
 sequencing about 2,000 ~ 3,000 colonies in original cDNA  
 library, the abundant cDNAs were selected and amplified by  
 PCR reaction using vector region primer including T7  
 promoter as 5' primer and NidT14 as 3' primer. The PCR  
 products were used as template for synthesis of  
 biotinylated single stranded RNA by in vitro transcription  
 reaction. The synthesized RNA probes were hybridized with  
 antisense single stranded cDNAs prepared from original  
 library and incubated with avidin-gel. After removing  
 DNA-RNA hybrids by centrifuge, the substracted cDNA  
 libraries were constructed by transformation of the  
 remaining DNA into competent cells E. coli Top10F' with  
 electroporation method."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 437;  
 Best Local Similarity 99.2%; Pred. No. 1.1e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGACCATAGACCAACAGGG 60  
 Db 174 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGACCATAGACCAACAGGG 233  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAGAGAGAG 120  
 Db 234 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAGAGAGAG 293  
 QY 121 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 294 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 353  
 QY 181 GCTAACATTTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 354 GCTAACATTTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACT 413  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 414 CAGGCCACCAACCGAT 428

## RESULT 2

BF819626/c  
 LOCUS BF819626 484 bp mRNA linear EST 13-JAN-2001  
 DEFINITION MRI-R10028-101100-002-b04 RT0028 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF819626  
 VERSION BF819626.1 GI:12157598  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiridae; Homo.  
 1 (bases 1 to 484)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Balá, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,

## TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR1&t2=MR1-RT0028-  
 101100-002-b04&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 484.

## FEATURES

. Location/Qualifiers  
 source

1..484

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="RT0028"

/note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 10; Length 484;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGACCATAGACCAACAGGG 60  
 Db 465 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGACCATAGACCAACAGGG 406  
 QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAGAGAGAG 120  
 Db 405 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAGAGAGAG 346  
 QY 121 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 345 ACCGTCCTGGCATCTGGAGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 286  
 QY 181 GCTAACATTTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 285 GCTAACATTTGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACT 226  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 225 CAGGCCACCAACCGAT 211

## RESULT 3

BF694247  
 LOCUS BF694247 491 bp mRNA linear EST 28-FEB-2002  
 DEFINITION UI-E-C11-af0-1-16-0-UI-r1 UI-E-C11 Homo sapiens cDNA clone  
 ACCESSION BF694247  
 VERSION BF694247.1 GI:19007505  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 491)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 PUBMED  
 8889548  
 COMMENT  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@iowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse  
 Location/Qualifiers  
 1..491  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-B-C11-af0-1-16-0-UI"  
 /tissue\_type="RPE and Choroid"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-B-C11"  
 /notes="Organ: eye; Vector: pVT3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; site 2: Not I;  
 UI-B-C11 is a normalized cDNA library containing the  
 following tissue(s): RPE and Choroid. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pVT3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is ACTTA.  
 This library was created for the program, Gene Discovery  
 in the Visual System, supported by National Eye Institute  
 (NEI)."

RESULT 4  
 LOCUS  
 CD698819  
 DEFINITION  
 ACCESSION  
 CD698819  
 VERSION  
 CD698819.1  
 GI:32227504  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 514)  
 AUTHORS  
 Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and  
 Zeng,X.-X.  
 TITLE  
 Transcriptional Gene Expression Profile of Human Nasopharynx  
 Unpublished (2003)  
 JOURNAL  
 COMMENT  
 Contact: YiXin Zeng  
 Cancer Center  
 Sun Yat-sen University  
 651 Dongfeng Road East, Guangzhou 510060, China  
 Tel: 86-1380-9770-743  
 Fax: 86-20-8775-4506  
 Email: yxzeng@gzsums.edu.cn.  
 Location/Qualifiers  
 1..514  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA  
 library from Southern Chinese"

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 14; Length 514;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTCTCAACTTATCCCGCTTTGTACAGCATAGACCAACAGGG 60  
 Db 208 ATCAAGCGGACCATGTCTCAACTTATCCCGCTTTGTACAGCATAGACCAACAGGG 267  
 QY 61 GAGTTTATGTTTGAATTTGATGAGATCAGATCAGATGTTCTATGTGGATCTGACAGAGAGAG 120  
 Db 268 GAGTTTATGTTTGAATTTGATGAGATCAGATGTTCTATGTGGATCTGACAGAGAGAG 327  
 QY 121 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 328 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG 387  
 QY 181 GCTAACATTCGTATATTCGAACAACAACCTTGAATACCTTCATCCAGCGTCC-ACCACACT 239  
 Db 388 GCTAACATTCGTATATTCGAACAACAACCTTGAATACCTTCATCCAGCGTCC-ACCACACT 447  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 448 CAGGCCACCAACCGAT 462

RESULT 5  
 LOCUS  
 BM769742  
 DEFINITION  
 K-EST0053050 S14K402 Homo sapiens cDNA clone S14K402-25-B02 5',  
 mRNA sequence.  
 ACCESSION  
 BM769742  
 VERSION  
 BM769742.1  
 GI:19099357  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 515)  
 AUTHORS Kim.N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 25 row: B column: 02  
 High quality sequence stop: 515.

FEATURES  
 source  
 1..515  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S14K402-25-B02"  
 /cell\_line="K402"  
 /lab\_host="Top10F"  
 /clone\_lib="S14K402"  
 /note="Organ: Stomach; Vector: pTZ19REP1; Site 1: EcoRI;  
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."

ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 515;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGGGGACCATGTCTCACTTATGCCCGCTTTGTACAGACGATAGACCAACAGGG 60  
 Db 151 ATCAAGGGGACCATGTCTCACTTATGCCCGCTTTGTACAGACGATAGACCAACAGGG 210

QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 Db 211 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 270

QY 121 ACCGTCGCACTCTGGAGAGTTTGGCCACGCTTTCTTTAGGCTCAGGGGGGCTG 180  
 Db 271 ACCGTCGCACTCTGGAGAGTTTGGCCACGCTTTCTTTAGGCTCAGGGGGGCTG 330

QY 181 GCTACATGCTATATTGAACAACACTTGAATACCTTGAATCCAGCGCTCC-ACCACACT 239  
 Db 331 GCTACATGCTATATTGAACAACACTTGAATACCTTGAATCCAGCGCTCCACCACT 390

QY 240 CAGGCCACCAACCGAT 254  
 Db 391 CAGGCCACCAACCGAT 405

RESULT 6  
 BG5411135  
 LOCUS  
 DEFINITION 602569911F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4694379 5',  
 mRNA sequence.  
 ACCESSION BG5411135

VERSION BG5411135.1 GI:13533368  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 519)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L10M1518 row: e column: 04  
 High quality sequence stop: 513.  
 Location/Qualifiers  
 1..519  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4694379"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgcccgccc); Site 2: SfiI (ggccattggcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGGCCATATATGCC-3', and 3' adaptor sequence:  
 5'-ATCTAGAGCGGAGGCGCGCATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC library."

ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 519;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGGGGACCATGTCTCACTTATGCCCGCTTTGTACAGACGATAGACCAACAGGG 60  
 Db 174 ATCAAGGGGACCATGTCTCACTTATGCCCGCTTTGTACAGACGATAGACCAACAGGG 233

QY 61 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 120  
 Db 234 GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGGACAAGAGGAG 293

QY 121 ACCGTCGCACTCTGGAGAGTTTGGCCACGCTTTCTTTAGGCTCAGGGGGGCTG 180  
 Db 294 ACCGTCGCACTCTGGAGAGTTTGGCCACGCTTTCTTTAGGCTCAGGGGGGCTG 353

QY 181 GCTACATGCTATATTGAACAACACTTGAATACCTTGAATCCAGCGCTTC-ACCACACT 239  
 Db 354 GCTACATGCTATATTGAACAACACTTGAATACCTTGAATCCAGCGCTTCACCACT 413

QY 240 CAGGCCACCAACCGAT 254  
 Db 414 CAGGCCACCAACCGAT 428

RESULT 7  
 CD695435  
 LOCUS  
 DEFINITION EST11958 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
 ACCESSION CD695435  
 VERSION CD695435.1 GI:32220983  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 540)  
 AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
 TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: YiXin Zeng  
 Cancer Center  
 Sun Yat-sen University  
 651 Dongfeng Road East, Guangzhou 510060, China  
 Tel: 86-1380-9770-743  
 Fax: 86-20-8775-4506  
 Email: yxzeng@gzsums.edu.cn

FEATURES  
 source  
 1..540  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 14; Length 540;  
 Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60  
 DB 181 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 240  
 QY 61 GAGTTTATGTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGGACAGAGGAG 120  
 DB 241 GAGTTTATGTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGGACAGAGGAG 300  
 QY 121 ACCGTCCTGGCATCTGGAGGAGTTGGCAAGCTTTTCCTTTGAGGCTCAGGGGGCTG 180  
 DB 301 ACCGTCCTGGCATCTGGAGGAGTTGGCAAGCTTTTCCTTTGAGGCTCAGGGGGCTG 360  
 QY 181 GCTAACCTTGGTATTTGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACACT 239  
 DB 361 GCTAACCTTGGTATTTGACACAACTTGAATACCTTGATCCAGGCTTCCACACACT 420  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 421 CAGGCCACCAACCGAT 435

RESULT 8  
 BU783326  
 LOCUS  
 DEFINITION in02f04.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123391  
 5' similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.  
 BU783326  
 ACCESSION BU783326.1 GI:23827397  
 VERSION BU783326.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 545)  
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,C., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
 TITLE Endocrine Pancreas Consortium

JOURNAL COMMENT  
 Unpublished (2000)  
 Other ESTs: in02f04.x1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 431.  
 Location/Qualifiers  
 1..545  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="human nasopharynx"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 13; Length 545;  
 Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60  
 DB 131 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 190  
 QY 61 GAGTTTATGTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGGACAGAGGAG 120  
 DB 191 GAGTTTATGTTGAATTGATGAGATGAGATGTTCTATGTGGATCTGGACAGAGGAG 250  
 QY 121 ACCGTCCTGGCATCTGGAGGAGTTGGCAAGCTTTTCCTTTGAGGCTCAGGGGGCTG 180  
 DB 251 ACCGTCCTGGCATCTGGAGGAGTTGGCAAGCTTTTCCTTTGAGGCTCAGGGGGCTG 310  
 QY 181 GCTAACCTTGGTATTTGACACAACTTGAATACCTTGATCCAGGCTTCC-ACCACACT 239  
 DB 311 GCTAACCTTGGTATTTGACACAACTTGAATACCTTGATCCAGGCTTCCACACACT 370  
 QY 240 CAGGCCACCAACCGAT 254  
 DB 371 CAGGCCACCAACCGAT 385

RESULT 9  
 BG535978  
 LOCUS  
 DEFINITION 602564150F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4688604 5', mRNA  
 BG535978  
 ACCESSION BG535978  
 VERSION BG535978.1 GI:13527523  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 557)

NIH-MGC <http://imgc.nci.nih.gov/>. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [sgaps@email.nih.gov](mailto:sgaps@email.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Plate: LNCM1503 row: d column: 13  
High quality sequence stop: 557.

**FEATURES**  
source  
1..557  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4688604"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 77"  
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGAGCGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

**ORIGIN**  
Query Match 95.0%; Score 241.4; DB 12; Length 557;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 233; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
197 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 256  
Db  
QY 61 GAGCTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACGAGAGGAG 120  
257 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACGAGAGGAG 316  
Db  
QY 121 ACCGTCGTCATCTGAGGAGTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180  
317 ACCGTCGTCATCTGAGGAGTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 376  
Db  
QY 181 GCTAACATTGCTATATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239  
377 GCTAACATTGCTATATGAACAACTTGAATACCTTGATCCAGCGTTCCACCAACAT 436  
QY 240 CAGGCCACCAACCAT 254  
437 CAGGCCACCAACCAT 451  
Db

**RESULT 10**  
CB265399  
LOCUS  
DEFINITION  
1004304 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION  
CB265399  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 563)

Yang, R.-Z., Schuldiner, A. and Gong, D.-W. EST analysis of human adipose gene expression Unpublished (2002)  
Contact: Gong Da-Wei  
Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St, HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622  
Email: [dgong@medicine.umaryland.edu](mailto:dgong@medicine.umaryland.edu)  
PCR Primers  
FORWARD: CTCGGGAAGCGCGCATTTGTGTGGT  
BACKWARD: AATAGACTACATATAGGGGGAATTGG  
Seq primer: GTTGGTACCCGGGAATTC.  
Location/Qualifiers  
1..563  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="Male and female"  
/tissue\_type="Adipose"  
/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
/note="Vector: lambdaTriplex"

**ORIGIN**  
Query Match 95.0%; Score 241.4; DB 14; Length 563;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 60  
155 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTTGTACAGCGCATAGACCAACAGGG 214  
Db  
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACGAGAGGAG 120  
215 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACGAGAGGAG 274  
Db  
QY 121 ACCGTCGTCATCTGAGGAGTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180  
275 ACCGTCGTCATCTGAGGAGTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 334  
QY 181 GCTAACATTGCTATATGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239  
335 GCTAACATTGCTATATGAACAACTTGAATACCTTGATCCAGCGTTCCACCAACAT 394  
QY 240 CAGGCCACCAACCAT 254  
395 CAGGCCACCAACCAT 409  
Db

**RESULT 11**  
CA942442  
LOCUS  
DEFINITION  
ir56a06.y1 Hr65 islet Homo sapiens cDNA clone IMAGE:6549227 5', similar to SW-620 HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.  
ACCESSION  
CA942442  
VERSION  
CA942442.1 GI:27430922  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 571)

REFERENCE  
AUTHORS  
Melton, D., Brown, J., Kenny, G., Permutt, A., Lee, C., Kastner, K., Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T., Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium

JOURNAL  
COMMENT  
Unpublished (2000)  
Other ESTs: ir58a06.xl  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@chp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 487.

## FEATURES

source

Location/Qualifiers  
1..571  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6549227"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming. ~1kb. 5'  
Size: selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110. E-mail: hinoue@mgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 14; Length 571;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCACATTATGCGCGCTTTGTACAGCGATAGACCAACAGGG 60  
DB 155 ATCAAGCGGACCATGTGTCACATTATGCGCGCTTTGTACAGCGATAGACCAACAGGG 214  
QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTTGACAGAGAGAG 120  
DB 215 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTTGACAGAGAGAG 274  
QY 121 ACCGTCCTGGCATCTGAGGAGTTTGCCCAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
DB 275 ACCGTCCTGGCATCTGAGGAGTTTGCCCAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 334  
QY 181 GCTAACATTGCTATATGAAACAACTTGAATACCTTGATCAGCGTTCC-ACCACT 239  
DB 335 GCTAACATTGCTATATGAAACAACTTGAATACCTTGATCAGCGTTCC-ACCACT 394  
QY 240 CAGGCGACCAACCGAT 254  
DB 395 CAGGCGACCAACCGAT 409

RESULT 12  
LOCUS BM831052  
DEFINITION K-EST0104833 S14K402s1 Homo sapiens cDNA clone S14K402s1-18-Cl2 5',  
mRNA sequence.  
ACCESSION BM831052  
VERSION BM831052.1 GI:19187461  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS

1 (bases 1 to 579)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)

## TITLE

JOURNAL

COMMENT

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 18 row: C column: 12  
High quality sequence stop: 579.

## FEATURES

source

Location/Qualifiers  
1..579  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S14K402s1-18-Cl2"  
/cell\_line="K402"  
/lab\_host="Top10P"  
/clone\_lib="S14K402s1"  
/notes="Organ: Stomach; Vector: pTZ18R21; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10P by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library. After analyzing and  
sequencing about 2,000 ~ 3,000 colonies in original cDNA  
library, the abundant cDNAs were selected and amplified by  
PCR reaction using vector region primer including T7  
promoter as 5' primer and N(dT)14 as 3' primer. The PCR  
products were used as template for synthesis of  
biotinylated single stranded RNA by in vitro transcription  
reaction. The synthesized RNA probes were hybridized with  
antisense single stranded cDNAs prepared from original  
library and incubated with avidin-gel. After removing  
DNA-RNA hybrids by centrifuge, the subtracted cDNA  
libraries were constructed by transformation of the  
remaining DNA into competent cells E. coli Top10P with  
electroporation method."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 579;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCACATTATGCGCGCTTTGTACAGCGATAGACCAACAGGG 60  
DB 111 ATCAAGCGGACCATGTGTCACATTATGCGCGCTTTGTACAGCGATAGACCAACAGGG 170  
QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTTGACAGAGAGAG 120  
DB 171 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTTGACAGAGAGAG 230  
QY 121 ACCGTCCTGGCATCTGAGGAGTTTGCCCAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
DB 231 ACCGTCCTGGCATCTGAGGAGTTTGCCCAAGCCCTTTTCCTTTGAGGCTCAGGCGGGCTG 290  
QY 181 GCTAACATTGCTATATGAAACAACTTGAATACCTTGATCAGCGTTCC-ACCACT 239

```

Db      291  GCTACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCCACACCACT 350
QY      240  CAGGCCACCAACCGAT 254
Db      351  CAGGCCACCAACCGAT 365

RESULT 13
BM737984
LOCUS      K-EST0002067 S1SNUS Homo sapiens cDNA clone S1SNUS-25-B01 5', mRNA
DEFINITION
ACCESSION  BM737984
VERSION     BM737984.1 GI:19059313
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 582)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 25 row: B column: 01
High quality sequence stop: 582.

FEATURES
            source
            1..582
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S1SNUS-25-B01"
                /sex="F"
                /tissue_type="Ascites"
                /cell_type="Lymphoblast-like"
                /cell_line="GNU-5"
                /lab_host="Top10P"
                /clone_lib="S1SNUS"
                /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tobacco acid pyrophosphatase (TAP). The decapped
                intact mRNA was ligated with DNA-RNA linker including EcoR
                I site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized from oligo dT-selected mRNA by
                priming with dT-tailed vector. The dT-tailed vector was
                adjusted to have about 60nt. The cDNA vector was
                circularized with E. coli DNA ligase after digestion of
                EcoRI which site is also included in vector. An RNA strand
                converted to a DNA strand by Okayama-Berg method. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10P by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library."

ORIGIN
Query Match      95.0%; Score 241.4; DB 12; Length 582;
Best Local Similarity 99.2%; Pred. No. 1.3e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1  ATCAGGCGGACCATGTGTCACACTATCCCGGTTGTACAGACGATAGACCAACAGGG 60
Db      174  ATCAGGCGGACCATGTGTCACACTATCCCGGTTGTACAGACGATAGACCAACAGGG 233
QY      61  GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTCGACAGAAGGAG 120

```

```

Db      234  GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGACAGAAGGAG 293
QY      121  ACCGTCTGGCATCTGGAGGAGTTTGCCCAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
Db      294  ACCGTCTGGCATCTGGAGGAGTTTGCCCAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 353
QY      181  GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239
Db      354  GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCCACCACT 413
QY      240  CAGGCCACCAACCGAT 254
Db      414  CAGGCCACCAACCGAT 428

RESULT 14
AV706521
LOCUS      ADB Homo sapiens cDNA clone ADBAVE11 5', mRNA sequence.
DEFINITION
ACCESSION  AV706521
VERSION     AV706521.1 GI:10723800
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 585)
AUTHORS   Peng,Y., Song,H., Huang,Q., Huang,C., Gu,Y., Yang,Y., Gao,G.,
Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z.,
Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S.,
Zhong,M., Lu,G., Fu,R., Chen,J., Chen,Z. and Han,Z.
TITLE     Homo sapiens cDNA ADB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
            source
            1..585
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="ADBAVE11"
                /tissue_type="Adrenal gland"
                /dev_stage="Adult"
                /lab_host="SOLR"
                /clone_lib="ADB"
                /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                XhoI"

ORIGIN
Query Match      95.0%; Score 241.4; DB 9; Length 585;
Best Local Similarity 99.2%; Pred. No. 1.3e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1  ATCAAGCGGACCATGTGTCACACTATCCCGGTTGTACAGCGATAGACCAACAGGG 60
Db      225  ATCAAGCGGACCATGTGTCACACTATCCCGGTTGTACAGCGATAGACCAACAGGG 284
QY      61  GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGACAGAAGGAG 120
Db      285  GAGTTTATGTTTGAATTTGATGAAGATGAGATGTTCTATGTGGATCTGACAGAAGGAG 344
QY      121  ACCGTCTGGCATCTGGAGGAGTTTGCCCAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 180
Db      345  ACCGTCTGGCATCTGGAGGAGTTTGCCCAAGCTTTTCCCTTTGAGGCTCAGGGCGGGCTG 404
QY      181  GCTAACATTGCTATATTGAACACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239

```

Db	108	ATCAAGCGGACCATGTGTCACCTTATGCGCGTTGTACAGACCATAGACCAACAGGG	167	
QY	61	GAGTTTATCTTTGAATTTGATGAACATGAGATGTTCTATGTGGATCTGGACAAGAAGGAG	120	
Db	168	GAGTTTATGTTTGAATTTGATGAACATGAGATGTTCTATGTGGATCTGGACAAGAAGGAG	227	
QY	121	ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGGGGGCTG	180	
Db	228	ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCTTTTCCCTTTGAGGCTCAGGGGGGGCTG	287	
QY	181	GCTAACATTTGCTATATTAGAACACAACTTGATACCTTGATCCAGCGTTC-ACCACACT	239	
Db	288	GCTAACATTTGCTATATTAGAACACAACTTGATACCTTGATCCAGCGTTCACACACT	347	
QY	240	CAGGCCACCAACGAT	254	
Db	348	CAGGCCACCAACGAT	362	
RESULT 16				
BI9111442				
LOCUS	603063356F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212669 5', mRNA linear EST 16-OCT-2001			
DEFINITION	mRNA sequence.			
ACCESSION	BI9111442			
VERSION	BI9111442			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 586)			
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-rc@mail.nih.gov">cgapbs-rc@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. CDNA Library Prepared by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM1533 row: p column: 14 High quality sequence start: 18 High quality sequence stop: 585.			
FEATURES				
source				
1..589				
Location/Qualifiers				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="IMAGE:5212669"				
/tissue_type="leukocyte"				
/lab_host="DH10B"				
/clone_lib="NIH_MGC_118"				
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC library."				
ORIGIN				
Query Match 95.0%; Score 241.4; DB 12; Length 589;				
Best Local Similarity 99.2%; Pred. No. 1.3e-60;				
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;				
QY	1	ATCAAGCGGACCATGTGTCACCTTATGCGCGTTGTACAGACCATAGACCAACAGGG	60	

Db 166 ATCAAGCGGACCATGTGTCAACTATGCGCGGTTTGTACAGACGATAGACCAACAGGG 225

Qy 61 GAGTTATGTTTGAATTTGATCAAGATCAGATGTTCTATGTGGATCTGCAACAAGAGAG 120

Db 226 GAGTTATGTTTGAATTTGATCAAGATCAGATGTTCTATGTGGATCTGCAACAAGAGAG 285

Qy 121 ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTTCAGGCGGGCTG 180

Db 286 ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTTCAGGCGGGCTG 345

Qy 181 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCACACACT 239

Db 346 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCACACACT 405

Qy 240 CAGGCCACCAACCGAT 254

Db 406 CAGGCCACCAACCGAT 420

RESULT 17  
BM876262  
LOCUS  
DEFINITION i60c02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5635202  
3' similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
ANTIGEN, DP ALPHA CHAIN PRECURSOR ; mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BM876262  
BM876262.1 GI:19243928  
EST.  
Homo sapiens (human)  
Homo sapiens

Other\_ESTRs: i60c02.y1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioph.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 482.

Location/Qualifiers  
1. .592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/c\_one="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further

FEATURES  
SOURCE  
1. .592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/c\_one="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further

FEATURES  
SOURCE  
1. .592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/c\_one="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further

FEATURES  
SOURCE  
1. .592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/c\_one="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further

FEATURES  
SOURCE  
1. .592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/c\_one="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt  
Laboratory, Washington University School of Medicine, Box  
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
is a Washington University Pancreas EST project library. "

# ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 592;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 25; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGGACCATGTGTCAACTATGCGCGGTTTGTACAGACGATAGACCAACAGGG 60

Db 149 ATCAAGCGGACCATGTGTCAACTATGCGCGGTTTGTACAGACGATAGACCAACAGGG 208

Qy 61 GAGTTATGTTTGAATTTGATCAAGATCAGATGTTCTATGTGGATCTGCAACAAGAG 120

Db 209 GAGTTATGTTTGAATTTGATCAAGATCAGATGTTCTATGTGGATCTGCAACAAGAG 268

Qy 121 ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTTCAGGCGGGCTG 180

Db 269 ACCGTCTGGCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTTCAGGCGGGCTG 328

Qy 181 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCACACACT 239

Db 329 GCTAACATTCGTATATTGAACAACAACCTTGAATACCTTGATCCAGCGTTCACACACT 388

Qy 240 CAGGCCACCAACCGAT 254

Db 389 CAGGCCACCAACCGAT 403

RESULT 18  
BG926106  
LOCUS  
DEFINITION HNC50-1-B12.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BG926106  
BG926106.1 GI:14320629  
EST.  
Homo sapiens (human)  
Homo sapiens

Other\_ESTRs: i60c02.y1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioph.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 482.

Location/Qualifiers  
1. .594  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="cartilage"  
/lab\_host="E. coli DH10 B"  
/clone\_lib="HNC (Human Normal Cartilage)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Directional"

FEATURES  
SOURCE  
1. .594  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="cartilage"  
/lab\_host="E. coli DH10 B"  
/clone\_lib="HNC (Human Normal Cartilage)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Directional"

FEATURES  
SOURCE  
1. .594  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="cartilage"  
/lab\_host="E. coli DH10 B"  
/clone\_lib="HNC (Human Normal Cartilage)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Directional"

FEATURES  
SOURCE  
1. .594  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="cartilage"  
/lab\_host="E. coli DH10 B"  
/clone\_lib="HNC (Human Normal Cartilage)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Directional"

FEATURES  
SOURCE  
1. .594  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="cartilage"  
/lab\_host="E. coli DH10 B"  
/clone\_lib="HNC (Human Normal Cartilage)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Directional"

```

Query Match      95.0%; Score 241.4; DB 12; Length 594;
Best Local Similarity 99.2%; Pred. No. 1.3e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 60
Db 307 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 366
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
Db 367 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 426
QY 121 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 180
Db 427 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 486
QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 487 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 546
QY 240 CAGGCCACCAACCGAT 254
Db 547 CAGGCCACCAACCGAT 561

RESULT 19
LOCUS CD693703 595 bp mRNA linear EST 25-JUN-2003
DEFINITION EST10226 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD693703
VERSION CD693703.1 GI:32217621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.

FEATURES
source
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/notes="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match      95.0%; Score 241.4; DB 14; Length 595;
Best Local Similarity 99.2%; Pred. No. 1.3e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 60
Db 179 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 238
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
Db 239 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 298
QY 121 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 180
Db 121 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 347
QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 348 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 407
QY 240 CAGGCCACCAACCGAT 254
Db 408 CAGGCCACCAACCGAT 422

RESULT 21
LOCUS BQ270752 599 bp mRNA linear EST 15-JUL-2003

```

```

Db 299 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 358
QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 359 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 418
QY 240 CAGGCCACCAACCGAT 254
Db 419 CAGGCCACCAACCGAT 433

RESULT 20
LOCUS CD684394 597 bp mRNA linear EST 25-JUN-2003
DEFINITION EST914 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD684394
VERSION CD684394.1 GI:32199345
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.

FEATURES
source
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/notes="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match      95.0%; Score 241.4; DB 14; Length 597;
Best Local Similarity 99.2%; Pred. No. 1.3e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 60
Db 168 ATCAAGCGGACCATGTGTCAACTTATGCGCGGTTTGTACAGAGCATAGACCAACAGGG 227
QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
Db 228 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 287
QY 121 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 180
Db 288 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAAAGCCCTTTCTCTTTGAGGCTCAGGGGGGCTG 347
QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239
Db 348 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 407
QY 240 CAGGCCACCAACCGAT 254
Db 408 CAGGCCACCAACCGAT 422

RESULT 21
LOCUS BQ270752 599 bp mRNA linear EST 15-JUL-2003

```

ik04d01.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:5779993  
5' similar to SW:HA2Q.HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

DEFINITION BQ270752.1 GI:20495818  
ACCESSION BQ270752.1  
VERSION BQ270752.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 599)  
REFERENCE Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,  
Williams, T., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other ESTs: ik04d01.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Seq primer: -40sp from Gibco  
High quality sequence stop: 430.

FEATURES  
source  
1..599  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5779993"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1:  
xhoI; Site 2: EcoRI; Constructed with lambda ZAP1 system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further  
information on this library (Metabolism Division, Permutt  
Laboratory, Washington University School of Medicine, Box  
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
is a Washington University Pancreas EST project library."

ORIGIN  
Query Match 95.0%; Score 241.4; DB 13; Length 599;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 60  
108 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 167  
61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTTATGTCGTCGACAGAGGAG 123  
168 GAGTTTATGTTGAATTGATGAAGATGAGATGTTTATGTCGTCGACAGAGGAG 227  
121 ACCGTCGTCGATCGGAGGAGTTGGCCAAAGCCCTTTTCCTTCAGGCTCAGGCGGCTG 180  
228 ACCGTCGTCGATCGGAGGAGTTGGCCAAAGCCCTTTTCCTTCAGGCTCAGGCGGCTG 287  
181 GCTAACATTCGCTATTTGAACAACAACTTGATACCTTGATCCAGCGCTTCC-ACCACACT 239

Db 288 GCTACACATTCGCTATTTGAACAACAACTTGATACCTTGATCCAGCGTTCCACCACT 347  
Qy 240 CAGGCCACCAACCGAT 254  
Db 348 CAGGCCACCAACCGAT 362

RESULT 22  
LOCUS BF725429  
DEFINITION Bx15h06.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo  
sapiens cDNA clone bx15h06 5', mRNA sequence.  
ACCESSION BF725429  
VERSION BF725429.1 GI:12041348  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 601)  
REFERENCE Wistow, G.J., Bernstein, S., Behal, A. and Smith, D.  
NBIANK: EST analysis and bioinformatics for ocular genomics  
Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 15 row: h column: 06  
Seq primer: M13RPL reverse primer (ABI).  
Location/Qualifiers  
1..601  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="bx15h06"  
/tissue\_type="Iris"  
/dev\_stage="Adult"  
/lab\_host="EMDH103"  
/clone\_lib="Human Iris cDNA (Un-normalized, unamplified):  
BX"  
/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris  
tissue was pooled from 10 individuals ranging in age from  
4-80 years and RNA was extracted. From this pooled sample  
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A  
directionally cloned cDNA library in the pCMVSPORT6 vector  
was constructed at Life Technologies, essentially  
following the protocols of the Superscript Plasmid System  
full details of which are contained in the manufacturer's  
instruction manual (http://www.lifetech.com/). First  
strand synthesis was carried out using a Not I  
primer-adaptor  
[5'-PGACTAGTCTAGATCGGACGGCGGCC(T)15-3']. Not I/blunt  
end inserts were cloned into the Not I/EcoR V sites in the  
vector. EST analysis was performed on the unamplified  
library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN  
Query Match 95.0%; Score 241.4; DB 10; Length 601;  
Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 60  
98 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 157  
61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTTATGTCGTCGACAGAGGAG 120  
Db 158 GAGTTTATGTTGAATTGATGAAGATGAGATGTTTATGTCGTCGACAGAGGAG 217

QY 121 ACCGCTGGCATCTGGAGGAGTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGCGGGCTG 180  
 Db 218 ACCGCTGGCATCTGGAGGAGTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGCGGGCTG 277

QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 278 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACACT 337

QY 240 CAGGCCACCAACCGAT 254  
 Db 338 CAGGCCACCAACCGAT 352

RESULT 23  
 BU783579

LOCUS BU783579 602 bp mRNA linear EST 11-OCT-2002  
 DEFINITION ir06c10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123643  
 5' similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
 ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

ACCESSION BU783579  
 VERSION BU783579.1 GI:23827903  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 602)

REFERENCE Melton D., Brown J., Keny G., Permutt A., Lee C., Kaestner K.,  
 Lemiska I., Seearce M., Brestelli J., Gradwohl G., Clifton S.,  
 Hallier L., Marra M., Pape D., Wylie T., Martin J., Blistain A.,  
 Schmitt A., Theising B., Ritter E., Ronko I., Bennett J.,  
 Cardenas M., Gibbons M., McCann R., Cole R., Tsagarisvili R.,  
 Williams T., Jackson V. and Bowers V.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)

TITLE Other ESTs: ind0c10.x1  
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 COMMENT Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biochem.harvard.edu  
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
 University Genome Sequencing Center For information on obtaining a  
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 431.

FEATURES  
 source  
 1..602  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6123643"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site:1:  
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/est/lambda\_protocol.shtml).  
 Please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permutt  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 13; Length 602;  
 Best Local Similarity 99.12%; Pred. No. 1.3e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Best Local Similarity 99.2%; Pred. No. 1.3e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCTAGACCAACAGGG 60  
 Db 144 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCTAGACCAACAGGG 203

QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATTTCTATGTGATCTGGACAAGAGGAG 120  
 Db 204 GAGTTTATGTTGAATTTGATGAAGATGAGATTTCTATGTGATCTGGACAAGAGGAG 263

QY 121 ACCGCTGGCATCTGGAGGAGTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGCGGGCTG 180  
 Db 264 ACCGCTGGCATCTGGAGGAGTTGGCCAGCCCTTTCCCTTTGAGGCTCAGGCGGGCTG 323

QY 181 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 324 GCTAACATTGCTATTTGAACAACTTGAATACCTTGATCCAGCGTTCCACCACACT 383

QY 240 CAGGCCACCAACCGAT 254  
 Db 384 CAGGCCACCAACCGAT 398

RESULT 24  
 CA405960

LOCUS CA405960 610 bp mRNA linear EST 07-NOV-2002  
 DEFINITION 1002109 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 cDNA 5', mRNA sequence.

ACCESSION CA405960  
 VERSION CA405960.1 GI:24770831  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 610)

REFERENCE Yang R.-Z., Shuldiner A. and Gong D.-W.  
 JOURNAL EST analysis of human adipose gene expression  
 Unpublished (2002)  
 COMMENT Contact: Gong Da-Wei  
 Division of Endocrinology, Diabetes and Nutrition  
 University of Maryland  
 660 Redwood St, HH497, Baltimore, MD 21201, USA  
 Tel: 410 706 1672  
 Fax: 410 706 1622  
 Email: dgong@medicine.umaryland.edu  
 PCR Primers  
 FORWARD: CTCGGGAAGCGCGCCATGTGTGGT  
 BACKWARD: AATACGACTCATATAGGCGCAATTGG  
 Seq primer: GTTGGTACCCGGAATTC.

FEATURES  
 source  
 1..610  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /sex="Male and Female"  
 /tissue\_type="Adipose"  
 /clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
 /note="Vector: lambdaTriplex"

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 14; Length 610;  
 Best Local Similarity 99.12%; Pred. No. 1.3e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCTAGACCAACAGGG 60  
 Db 144 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCTAGACCAACAGGG 203

QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATTTCTATGTGATCTGGACAAGAGGAG 120  
 Db 204 GAGTTTATGTTGAATTTGATGAAGATGAGATTTCTATGTGATCTGGACAAGAGGAG 263



QY 121 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 264 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 323  
 QY 181 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 324 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 383  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 384 CAGGCCACCAACCGAT 398

RESULT 25  
 BG545380  
 LOCUS 602572611F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4700906 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG545380  
 VERSION BG545380.1 GI:13543873  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 626)  
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC);  
 TITLE Unpublished (1999)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1535 row: e column: 03  
 High quality sequence stop: 626.

FEATURES  
 source  
 1..626  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4700906"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC 77"  
 /notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGGCCATATGGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGCGGCACATG-dt (30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 12; Length 626;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 1 ATCAAGCGGACCATGTGTCACTTATGCGCGCTTGTACAGACGCATAGACCAACAGGG 60  
 161 ATCAAGCGGACCATGTGTCACTTATGCGCGCTTGTACAGACGCATAGACCAACAGGG 220  
 QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATCTTATGTGGATCTGGACAAGAGGAG 120  
 Db 221 GAGTTTATGTTGAATTTGATGAAGATGAGATCTTATGTGGATCTGGACAAGAGGAG 280

QY 121 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 281 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 340  
 QY 181 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 341 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 400  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 401 CAGGCCACCAACCGAT 415

RESULT 26  
 CB267128  
 LOCUS 1006034 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 DEFINITION cDNA 5', mRNA sequence.  
 ACCESSION CB267128  
 VERSION CB267128.1 GI:28441714  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 627)  
 REFERENCE Yang,R.-Z., Shuldiner,A. and Gong,D.-W.  
 AUTHORS EST analysis of human adipose gene expression  
 TITLE Unpublished (2002)  
 JOURNAL  
 COMMENT Contact: Gong Da-Wei  
 Division of Endocrinology, Diabetes and Nutrition  
 University of Maryland  
 660 Redwood St, HH497, Baltimore, MD 21201, USA  
 Tel: 410 706 1672  
 Fax: 410 706 1622  
 Email: dgong@medicine.umaryland.edu  
 PCR Primers  
 FORWARD: CACGGGAGCGGCGCATTTGTGTGTGGT  
 BACKWARD: ATACGACTACTATAGGCGGATTTGG  
 Seq primer: GTTGTACCGGGATTC.

FEATURES  
 source  
 1..627  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /sex="Male and Female"  
 /tissue\_type="Adipose"  
 /clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
 /note="Vector: lambdaTriplex"

ORIGIN  
 Query Match 95.0%; Score 241.4; DB 14; Length 627;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 1 ATCAAGCGGACCATGTGTCACTTATGCGCGCTTGTACAGACGCATAGACCAACAGGG 60  
 99 ATCAAGCGGACCATGTGTCACTTATGCGCGCTTGTACAGACGCATAGACCAACAGGG 158

QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATCTTATGTGGATCTGGACAAGAGGAG 120  
 Db 159 GAGTTTATGTTGAATTTGATGAAGATGAGATCTTATGTGGATCTGGACAAGAGGAG 218  
 QY 121 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 180  
 Db 219 ACCGCTCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCCCTTTGAGGCTCAGGGCGGGCTG 278

QY 181 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 239  
 Db 279 GCTAACATTCCTATATTGAACAACAACCTTGGAATACCTTGATCCAGCGTTCC-ACCACACT 338  
 QY 240 CAGGCCACCAACCGAT 254

```

Db          339 CAGGCCACCAACGAT 353

RESULT 27
BG709234
LOCUS    602674620F1 NIH_MGC_96 635 bp mRNA linear EST 07-MAY-2001
DEFINITION
602674620F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797340 5',
mRNA sequence.
ACCESSION BG709234
VERSION   BG709234.1 GI:13987366
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 635)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLNL0683 row: 9 column: 05
          High quality sequence stop: 635.
FEATURES             source
     1..635
     /location/Qualifiers
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:4797340"
     /tissue_type="hypothalamus"
     /lab_host="DH10B"
     /clone_lib="NIH_MGC_96"
     /note="Organ: brain; Vector: pBluescriptR (modified
     pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
     (Gcgag); Oligo-dT primed using primer
     5'-TTTTTTTTTTTNN-3', size-selected for average
     insert size 2.3 kb and normalized to ROT 5. This is a
     primary library enriched for full-length clones and
     constructed using the Cap-trapper method (Carninci, in
     preparation) Library constructed by M. Brownstein
     (NIH/NHGRI, National Institutes of Health). Note: this is
     a NIH_MGC Library."

Db          339 CAGGCCACCAACGAT 406

RESULT 28
BM767502
LOCUS    643 bp mRNA linear EST 04-MAR-2002
DEFINITION
K-3870049979 S1SNUS52 Homo sapiens cDNA clone S1SNUS52-3-B03 5',
mRNA sequence.
ACCESSION BM767502
VERSION   BM767502.1 GI:19097117
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 643)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
          Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 3 row: B column: 03
          High quality sequence stop: 643.
FEATURES             source
     1..643
     /location/Qualifiers
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="S1SNUS52-3-B03"
     /sex="F"
     /tissue_type="Ascites"
     /cell_type="Lymphoblast-like"
     /cell_line="SNU-5"
     /lab_host="Top10F"
     /clone_lib="S1SNUS52"
     /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
     Site 2: NotI. The poly (A)+ RNA was dephosphorylated with
     bacterial alkaline phosphatase (BAP) and then decapped
     with tobacco acid pyrophosphatase (TAP). The decapped
     intact mRNA was ligated with DNA-RNA linker including EcoR
     I site by treatment of T4 RNA ligase and the first strand
     cDNA was synthesized from oligo dT-selected mRNA by
     priming with dT-tailed vector. The dT-tailed vector was
     adjusted to have about 60nt. The cDNA vector was
     circularized with E. coli DNA ligase after digestion of
     EcoRI which site is also included in vector. An RNA strand
     converted to a DNA strand by Okayama-Berg method. The
     obtained cDNA vectors were used for transposition of
     competent cells E. coli Top10F by electroporation method.
     The cDNA libraries constructed by this method are
     full-length enriched cDNA library. After analyzing and
     sequencing about 2,000 ~ 3,000 colonies in original cDNA
     library, the abundant cDNAs were selected and amplified by
     PCR reaction using vector region primer including 77
     promoter as 5' primer and N(dT)14 as 3' primer. The PCR
     products were used as template for synthesis of
     biotinylated single stranded RNA by in vitro transcription
     reaction. The synthesized RNA probes were hybridized with
     antisense single stranded cDNAs prepared from original
     library and incubated with avidin-gel. After removing
     DNA-RNA hybrids by centrifuge, the subtracted cDNA
     libraries were constructed by transformation of the
     remaining DNA into competent cells E. coli Top10F with
     electroporation method."

ORIGIN
Query Match          95.0%; Score 241.4; DB 12; Length 635;
Best Local Similarity 99.2%; Preq. No. 14e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGGACCAATGTCAACTATGCGCGTTGTACAGACGGCATAGACCAACAGGG 60
Db 152 ATCAAGCGGACCAATGTCAACTATGCGCGTTGTACAGACGGCATAGACCAACAGGG 211
Qy 61 GAGTTTATGTTGAATTTGATGAGATGAGATGTTCTATGTGATCTGGACAGAGGAG 120
Db 212 GAGTTTATGTTGAATTTGATGAGATGAGATGTTCTATGTGATCTGGACAGAGGAG 271
Qy 121 ACCGTCCTGGCACTGGAGGAGTTGGCAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 190
Db 272 ACCGTCCTGGCATCTGGAGGAGTTGGCAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 331
Qy 181 GCTAACATTCCTATTTGACCAACTTGAATACCTTGATCCAGGTTCC-ACCAACT 239
Db 332 GCTAACATTCCTATTTGACCAACTTGAATACCTTGATCCAGGTTCCACCAACT 391
Qy 240 CAGGCCACCAACGAT 254

```

Query Match	95.0%; Score 241.4; DB 12; Length 645;
Best Local Similarity	99.2%; Pred. NO. 1.4e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
QY	1 ATCAAGCGGACCATGTGCAACTTATGCCGGCTTTGTACAGACGCATAGCCAAACAGGG 60
Dd	156 ATCAAGCGGACCAATGTGCRACTTATGCCGGCTTTGTACAGACGCATAGCCAAACAGGG 215
QY	61 GAGTTTATGTTTGAAATTGATGAAGATGAGATGTTCTTATGTGATCTGGACAAGAAGGAG 120
Dd	216 GAGTTTATGTTTGAAATTGATGAAGATGAGATGTTCTTATGTGATCTGGACAAGAAGGAG 275
QY	121 ACCGTCCTGCATCTGAGAGAGTTTGCCAAGCCTTTTCCTTTGAGCCTCAGSCGGGCGCTG 180
Dd	276 ACCGTCGSCATCTGAGAGAGTTTGCCAAGCCTTTTCCTTTGAGCCTCAGSGCGGCGCTG 335
QY	181 GCTAACATNGCHATAATTGAACAACACTTTGAATACCTTGATCCAGCGTTC-ACCACACT 239
Dd	336 GCTAACATNGCTATAATTGAACAACACTTTGAATACCTTGATCCAGCGTTCACACCACT 395
QY	240 CAGGCCACCAACCGAT 254
Dd	396 CAGGCCACCAACCGAT 410
RESULT 30	
CD687077	
LOCUS	CD687077 659 bp mRNA linear EST 25-JUN-2003
DEFINITION	EST3598 human nasopharynx Homo sapiens cdna, mRNA sequence.
ACCESSION	CD687077
VERSION	CD687077.1 GI:32204592

```

RESULT 30
CD687077
LOCUS             659 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION       human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION        CD687077
VERSION          CD687077.1  GI:32204592
KEYWORDS
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS           Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
                  Zeng,Y.-X.
REFERENCE         Transcriptional Gene Expression Profile of Human Nasopharynx
                  Unpublished (2003)
JOURNAL           Contact: YiXin Zeng
COMMENT           Cancer Center
                  Sun Yat-sen University
                  651 DongFeng Road East, Guangzhou 510060, China
                  Tel: 86-1380-9770-743
                  Fax: 86-20-8775-4506
                  Email: yxzeng@gzsums.edu.cn.
FEATURES          Location/Qualifiers
source            1..659
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /tissue_type="normal nasopharynx"
                  /clone_lib="human nasopharynx"
                  /note="ESTs generated from a normal nasopharynx cDNA
                  library from southern Chinese"
ORIGIN
Query Match      95.0%; Score 241.4; DB 14; Length 659;
Best Local Similarity 99.2%; Pred. No. 1.4e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1  ATCAAGCGGACCAATGTGTCAACTTATGCCCGTTTGTCACAGACGCATAGACCAACAGG 60
244 ATCAAGCGGACCAATGTGTCAACTTATGCCCGTTTGTCACAGACGCATAGACCAACAGG 303

61  GAGTTTATGTTTGAAATTGATGAAGATCAGATGTTCTATGTGGATCTCGACAAAGACGAG 120
304 GAGTTTATGTTTGAAATTGATGAAGATCAGATGTTCTATGTGGATCTCGACAAAGACGAG 363

121 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAAAGCCCTTTCTTTGAGGCTCAGGGCGGCTG 180

```

121 ACGTCTGGCATCTGGAGGAGTTGGCCAGCCTTTTCCTTGAGGCTCAGGCGGGCTG 180

Db 364 ACCGTCCTGGCATCTGGAGAGCTTTGGCCAGCCCTTTCTCTTGGAGCTCAGGCGGCGTG 423

Qy 181 GCTAACATTTGCTATATTGAAACAACTTGTGATACCTTGATCCAGCGTTCC-ACCACACT 239

Db 424 GCTAACATTTGCTATATTGAAACAACTTGTGATACCTTGATCCAGCGTTCCACACACT 483

Qy 240 CAGGCCACCAACCGAT 254

Db 484 CAGGCCACCAACCGAT 498

RESULT 31

LOCUS AL048565

DEFINITION DKFZp586P0824 x1 586 (synonym: hute1) Homo sapiens cDNA clone

ACCESSION AL048565

VERSION DKFZp586P0824 mRNA sequence.

KEYWORDS EST.

SOURCE AL048565.2 GI:5936568

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 667)

AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Duesterhoeft, et al.)

JOURNAL Unpublished (1999)

COMMENT On Apr 30, 1999 this sequence version replaced gi:4727705.

CONTACT: MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp586P0824) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: c.lone@rzpd.de.

Location/Qualifiers

FEATURES

source 1..667

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp586P0824"

/tissue\_type="uterus"

/dev\_stage="adult"

/lab\_host="DE10B"

/clone\_lib="586 (synonym: hute1)"

/note="Vector: pSport1; Site 1: NotI; Site 2: SalI/XbaI"

ORIGIN

Query Match 95.0%; Score 241.4; DB 9; Length 667;

Best Local Similarity 99.2%; Pred. No. 1.4e-60;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGCGACCATGTCACACTTATGCGCGCTTTGACAGACGATAGACCAACAGG 60

Db 101 ATCAAGCGCGACCATGTCACACTTATGCGCGCTTTGACAGACGATAGACCAACAGG 160

Qy 61 GAGTTTATGTTGAAATTGATGAAGATGAGATGTTCTATGTGATCTGCAAGAGAGAG 120

Db 161 GAGTTTATGTTGAAATTGATGAAGATGAGATGTTCTATGTGATCTGCAAGAGAGAG 220

Qy 121 ACCGTCGTCATCGGAGGAGTTTGCCCAACCTTTTCTTTCGAGCGTCAGGCGGCGTG 180

Db 221 ACCGTCGTCATCGGAGGAGTTTGCCCAACCTTTTCTTTCGAGCGTCAGGCGGCGTG 280

Qy 181 GCTAACATTTGCTATATTGAAACAACTTGTGATACCTTGATCCAGCGTTCC-ACCACACT 239

Db 281 GCTAACATTTGCTATATTGAAACAACTTGTGATACCTTGATCCAGCGTTCCACACACT 340

Qy 240 CAGGCCACCAACCGAT 254

Db 341 CAGGCCACCAACCGAT 355

RESULT 32

LOCUS BU071266

DEFINITION im41D05.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6037569

ACCESSION BU071266

VERSION BU071266.1 GI:22512448

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 669)

AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y. and Bowers, Y.

TITLE Endocrine Pancreas Consortium

JOURNAL Unpublished (2000)

COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco

High quality sequence stop: 474.

Location/Qualifiers

FEATURES

source 1..669

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6037569"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system (stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol

(http://genome.wustl.edu/est/lambda\_protocol.shtml).

Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

ORIGIN

Query Match 95.0%; Score 241.4; DB 13; Length 669;

Best Local Similarity 99.2%; Pred. No. 1.4e-60;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATCAAGCGCGACCATGTCACACTTATGCGCGCTTTGACAGACGATAGACCAACAGG 60

Db 142 ATCAAGCGCGACCATGTCACACTTATGCGCGCTTTGACAGACGATAGACCAACAGG 201

QY 61 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 120  
 |||||  
 Db 202 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 261  
 |||||  
 QY 121 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 180  
 |||||  
 Db 262 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 321  
 |||||  
 QY 181 GCTAACATTGCTATTAATGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239  
 |||||  
 Db 322 GCTAACATTGCTATTAATGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACAT 381  
 |||||  
 QY 240 CAGGCCACCAACCGAT 254  
 |||||  
 Db 382 CAGGCCACCAACCGAT 396  
 |||||

RESULT 33  
 BG570300  
 LOCUS 676 bp mRNA linear EST 10-APR-2001  
 DEFINITION 602590723F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4717613 5', mRNA sequence.  
 ACCESSION BG570300  
 VERSION BG570300.1 GI:13577953  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 676)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: MGC Genomics, Inc.  
 Clone distribution by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LICM1567 row: m column: 06  
 High quality sequence stop: 669.  
 Location/Qualifiers  
 1. .676  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4717613"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcattaggcc); Site 2: SfiI (ggcattaggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCAATTATGGC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGGAGCGGCGGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## FEATURES

source  
 1. .676  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4717613"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcattaggcc); Site 2: SfiI (ggcattaggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCAATTATGGC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGGAGCGGCGGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 676;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGAGCATTGTCAACTTATGCGCGGTTGTACAGACGATAGACCAACAGG 60  
 |||||  
 Db 190 ATCAAGCGGAGCATTGTCAACTTATGCGCGGTTGTACAGACGATAGACCAACAGG 249  
 |||||  
 QY 61 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 120  
 |||||

Db 250 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 309  
 |||||  
 QY 121 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 180  
 |||||  
 Db 310 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 369  
 |||||  
 QY 181 GCTAACATTGCTATTAATGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239  
 |||||  
 Db 370 GCTAACATTGCTATTAATGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACAT 429  
 |||||  
 QY 240 CAGGCCACCAACCGAT 254  
 |||||  
 Db 430 CAGGCCACCAACCGAT 444  
 |||||

## RESULT 34

AV759427  
 LOCUS 683 bp mRNA linear EST 19-OCT-2000  
 DEFINITION AV759427 MDS Homo sapiens cDNA clone MDSALC03 5', mRNA sequence.  
 ACCESSION AV759427  
 VERSION AV759427.1 GI:10917275  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 683)  
 AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, P., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z., and Chen, Z.  
 TITLE Homo sapiens cDNA MDS clones  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: [hanzg@chgc.sh.cn](mailto:hanzg@chgc.sh.cn)  
 This clone is available at CHGC in Shanghai.

## FEATURES

source  
 1. .683  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MDSALC03"  
 /tissue\_type="Bone marrow"  
 /cell\_type="CD34+ hematopoietic stem/progenitor cell"  
 /lab\_host="BM25.8"  
 /clone\_lib="MDS"  
 /note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI"

## ORIGIN

Query Match 95.0%; Score 241.4; DB 9; Length 683;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGGAGCATTGTCAACTTATGCGCGGTTGTACAGACGATAGACCAACAGG 60  
 |||||  
 Db 160 ATCAAGCGGAGCATTGTCAACTTATGCGCGGTTGTACAGACGATAGACCAACAGG 219  
 |||||  
 QY 61 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 120  
 |||||  
 Db 220 GAGTTTATGTTGTAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGG 279  
 |||||  
 QY 121 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 180  
 |||||  
 Db 280 ACCGTCCTGGATCTGGAGGAGTTTGGCAAGCCTTTTCCTTTGAGGCTCAGGGGGGCTG 339  
 |||||  
 QY 181 GCTAACATTGCTATTAATGAACAACAACCTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239  
 |||||

```

Db      340  GCTAACATTGCTATATTGACACACACTTGAATACCTTGATCCAGCGTTCCACACCACT 399
QY      240  CAGGCCACCAACCGAT 254
Db      400  CAGGCCACCAACCGAT 414

RESULT 35
BE973719
LOCUS   601680707F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:3951145 5',
DEFINITION mRNA sequence.
ACCESSION BE973719
VERSION   BE973719.1 GI:10587055
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 685)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: CLONTECH Laboratories, Inc.
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCMS18 row: m column: 02
          High quality sequence stop: 614.
FEATURES             source
     Location/Qualifiers
         1..685
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:3951145"
             /lab_host="DH10B (T1 phage-resistant)"
             /clone_lib="NIH_MGC_83"
             /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
             Site_1: SfiI (ggccgctcgcc); Site_2: SfiI
             (ggccattatggcc); 5' and 3' adaptors were used in cloning
             as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3'
             and 3' adaptor sequence:
             5'-ATTCTAGAGCGGAGCGGCGGACATG-dT(30)BN-3' (where B = A,
             C, or G and N = A, C, G, or T). Average insert size 1.4
             kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
             by PCR. This library was enriched for full-length clones
             and was constructed by Clontech Laboratories (Palo Alto,
             CA)."
ORIGIN
Query Match      95.0%; Score 241.4; DB 10; Length 685;
Best Local Similarity 99.2%; Pred. No. 1.4e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1  ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60
Db      54  ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 113
QY      61  GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
Db      114  GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 173
QY      121  ACCGTCGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGCGGGGCTG 180
Db      174  ACCGTCGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGCGGGGCTG 233
QY      181  GCTAACATTGCTATATTGACACACACTTGAATACCTTGATCCAGCGTTCCACACACT 239

```

```

Db      234  GCTAACATTGCTATATTGACACACACTTGAATACCTTGATCCAGCGTTCCACACCACT 293
QY      240  CAGGCCACCAACCGAT 254
Db      294  CAGGCCACCAACCGAT 308

RESULT 36
BE439689
LOCUS   HTM1-360F HTM1 Homo sapiens cDNA, mRNA linear EST 25-JUL-2000
DEFINITION HTM1-360F HTM1 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE439689
VERSION   BE439689.1 GI:9439171
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 724)
AUTHORS   Gonzalez, P., Epstein, D.L. and Borras, T.
TITLE     Characterization of gene expression in human trabecular meshwork
          using single-pass sequencing of 1060 clones
JOURNAL   Invest. Ophthalmol. Vis. Sci. (2000) In press
COMMENT   Contact: Pedro Gonzalez
          Department of Ophthalmology
          Duke University
          Duke Eye Center, Erwin Rd, Box 3802, Durham, NC 27710-3802, USA
          Tel: 919 681 4085
          Fax: 919 684 8983
          Email: pedro.gonzalez@duke.edu.
FEATURES             source
     Location/Qualifiers
         1..724
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /tissue_type="Trabecular meshwork"
             /clone_lib="HTM1"
ORIGIN
Query Match      95.0%; Score 241.4; DB 10; Length 724;
Best Local Similarity 99.2%; Pred. No. 1.5e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1  ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60
Db      170  ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 229
QY      61  GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 120
Db      230  GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAAGAGGAG 289
QY      121  ACCGTCGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGCGGGGCTG 180
Db      290  ACCGTCGGCATCTGGAGAGTTTGGCCAAAGCCTTTCTTTGAGGCTCAGGCGGGGCTG 349
QY      181  GCTAACATTGCTATATTGACACACACTTGAATACCTTGATCCAGCGTTCCACACACT 239
Db      350  GCTAACATTGCTATATTGACACACACTTGAATACCTTGATCCAGCGTTCCACACACT 409
QY      240  CAGGCCACCAACCGAT 254
Db      410  CAGGCCACCAACCGAT 424

RESULT 37
AV733676
LOCUS   AV733676 cda Homo sapiens cDNA clone cdABC04 5', mRNA linear EST 17-OCT-2000
DEFINITION AV733676
ACCESSION AV733676
VERSION   AV733676.1 GI:10851221
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 734)  
Yang, Y., Song, H., Peng, Y., Gu, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.

## TITLE

Homo sapiens cDNA cda clones

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex. 45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

## FEATURES

source

Location/Qualifiers

1..734

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="cdABC04"

/tissue\_type="pheochromocytoma"

/dev\_stage="Adult"

/lab\_host="BW25.8"

/clone\_lib="cda"

/note="Vector: pRiplex2; Site\_1: sfIIA; Site\_2: sfIIB"

## ORIGIN

Query Match 95.0%; Score 241.4; DB 9; Length 734;  
Best Local Similarity 99.2%; Pred. No. 1.5e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 ATCAAGCGCGACCATGTCTCAACTTATCGCGCGTTTGTACAGACGCTAGACCAACAGGG 60
Db 173 ATCAAGCGCGACCATGTCTCAACTTATCGCGCGTTTGTACAGACGCTAGACCAACAGGG 232

QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATCTGATCTGACACAGAGGAG 120
Db 233 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATCTGATCTGACACAGAGGAG 232

QY 121 ACCGTCGTCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTGGAGGCTCAGGCGGGCTG 180
Db 293 ACCGTCGTCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTGGAGGCTCAGGCGGGCTG 352

QY 181 GCTAACATTGCTATTGACACAACTTGAACCTTGAATACCTTGTATCCAGGTTCC-ACCACACT 239
Db 353 GCTAACATTGCTATTGACACAACTTGAACCTTGAATACCTTGTATCCAGGTTCCACCAACT 412

QY 240 CAGGCCACCAACCGAT 254
Db 413 CAGGCCACCAACCGAT 427

```

## RESULT 38

BI762063

LOCUS

603049235F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:519361 5',  
mRNA sequence.

## ACCESSION

BI762063

VERSION

BI762063.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 735)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM11473 row: e column: 10

High quality sequence stop: 734.

## FEATURES

source

Location/Qualifiers

1..735

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:519361"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_116"

/note="Organ: pCooled colon, kidney, stomach; Vector:

PCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA

source anonymous pool of 3 clones, age 26 yo male, 49 yo

female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is

oligo-dr primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.4 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

023. Note: this is a NIH\_MGC Library."

## ORIGIN

```

Query Match 95.0%; Score 241.4; DB 12; Length 735;
Best Local Similarity 99.2%; Pred. No. 1.5e-60;
Matches 233; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGCGACCATGTCTCAACTTATCGCGCGTTTGTACAGACGCTAGACCAACAGGG 60
Db 146 ATCAAGCGCGACCATGTCTCAACTTATCGCGCGTTTGTACAGACGCTAGACCAACAGGG 205

QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTTGATCTGCAAGAGGAG 120
Db 206 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTTGATCTGCAAGAGGAG 255

QY 121 ACCGTCGTCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTGGAGGCTCAGGCGGGCTG 180
Db 266 ACCGTCGTCATCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTGGAGGCTCAGGCGGGCTG 325

QY 181 GCTAACATTGCTATTGACACAACTTGAACCTTGAATACCTTGTATCCAGGTTCC-ACCACACT 239
Db 326 GCTAACATTGCTATTGACACAACTTGAACCTTGAATACCTTGTATCCAGGTTCCACCAACT 385

QY 240 CAGGCCACCAACCGAT 254
Db 386 CAGGCCACCAACCGAT 400

```

## RESULT 39

CD104670

LOCUS

DEFINITION

AGENCY: 14006610 NIH\_MGC\_186 Homo sapiens cDNA clone

IMAGE:30369178 5', mRNA sequence.

CD104670

ACCESSION

VERSION

CD104670.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 743)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: Clontech Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: NDCM156 row: c column: 11  
High quality sequence stop: 569.

FEATURES  
source  
1. .743  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30369178"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_186"

/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: Sfil (ggcattatggcc); Site 2: Sfil (ggcgctcgcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a pooled samples of tissues from skin, meninges, dura mater, pia mater and choroid plexus. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGCGGCATG-dT(30)BN-3' (where B = A, C, G or G and N = A, C, G, or T). Average insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library"

## ORIGIN

Query Match 95.0%; Score 241.4; DB 14; Length 743;  
Best Local Similarity 99.2%; Pred. No. 1.5e-60;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 60  
DB 182 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 241  
QY 61 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTCATCTGACACAGAGGAG 120  
DB 242 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTCATCTGACACAGAGGAG 301  
QY 121 ACCGTCCTGGCATCTGGAGGAGTTGGCCAGACCTTTCTCTTTCAGGCTCAGGCGGGCTG 180  
DB 302 ACCGTCCTGGCATCTGGAGGAGTTGGCCAGACCTTTCTCTTTCAGGCTCAGGCGGGCTG 361  
QY 181 GCTAACATGCTATATGACACAACTTGAATACCTTGATCAGCGGTTCC-ACCACACT 239  
DB 362 GCTAACATGCTATATGACACAACTTGAATACCTTGATCAGCGGTTCC-ACCACACT 421  
QY 240 CAGGCCACCAACCAT 254  
DB 422 CAGGCCACCAACCAT 436

## RESULT 40

BI597401

LOCUS

DEFINITION 603243611F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5286122 5', mRNA sequence.

ACCESSION BI597401

VERSION BI597401.1 GI:15490340

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 744)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

## JOURNAL COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1AM11723 row: e column: 03

High quality sequence stop: 744.

## FEATURES

source

1. .744  
Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5286122"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_96"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(ctcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTTIVN-3', size-selected for average

insert size 2.3 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 744;  
Best Local Similarity 99.2%; Pred. No. 1.5e-60;

Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 60  
DB 178 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTGTACAGACGCATAGACCAACAGGG 237  
QY 61 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTCATCTGACACAGAGGAG 120  
DB 238 GAGTTTATGTTTGAATTGATGAGATGAGATGTTCTATGTCATCTGACACAGAGGAG 297  
QY 121 ACCGTCCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTCTCTTTCAGGCTCAGGCGGGCTG 180  
DB 298 ACCGTCCTGGCATCTGGAGGAGTTGGCCAAAGCCTTTCTCTTTCAGGCTCAGGCGGGCTG 357  
QY 181 GCTAACATGCTATATGACACAACTTGAATACCTTGATCAGCGGTTCC-ACCACACT 239  
DB 358 GCTAACATGCTATATGACACAACTTGAATACCTTGATCAGCGGTTCC-ACCACACT 417  
QY 240 CAGGCCACCAACCAT 254  
DB 418 CAGGCCACCAACCAT 432

## RESULT 41

CB984378

LOCUS

DEFINITION AGNCOURT\_13460933 NIH\_MGC\_184 Homo sapiens cDNA clone

IMAGE:30327118 5', mRNA sequence.

ACCESSION CB984378

VERSION CB984378.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 745)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS



TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: NDCM135 row: j column: 23  
 High quality sequence stop: 503.

## FEATURES

Location/Qualifiers  
 1..745  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30327118"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_184"  
 /note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: SfiI (ggccattatggcc); Site\_2: SfiI (ggcgctggcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCAGAGCCGAGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 14; Length 745;  
 Best Local Similarity 99.2%; Pred. No. 1.5e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGCGACCATGTGTCACATTATGCGGTTGTACAGACGATAGACCAACAGG 60  
 Db 160 ATCAAGCGCGACCATGTGTCACATTATGCGGTTGTACAGACGATAGACCAACAGG 219  
 QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGACAGAAGAGAG 120  
 Db 220 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGACAGAAGAGAG 279  
 QY 121 ACCGTCTGCATCTGGAGAGTTTGCCAAAGCTTTTCTTTGAGCTCAGGCGGGCTG 180  
 Db 280 ACCGTCTGCATCTGGAGAGTTTGCCAAAGCTTTTCTTTGAGCTCAGGCGGGCTG 339  
 QY 181 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCTTCCAAACACACT 239  
 Db 340 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCTTCCAAACACACT 399  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 400 CAGGCCACCAACCGAT 414

## RESULT 42

CB985816  
 LOCUS CB985816  
 DEFINITION AGENCOURT\_13458754 NIH\_MGC\_184 Homo sapiens cDNA clone  
 IMAGE:30326083 5', mRNA sequence.  
 ACCESSION CB985816  
 VERSION CB985816.1 GI:30280340  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1. (bases 1 to 758)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: NDCM132 row: o column: 20  
 High quality sequence stop: 471.

## FEATURES

Location/Qualifiers  
 1..758  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30326083"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_184"  
 /note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: SfiI (ggccattatggcc); Site\_2: SfiI (ggcgctggcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCAGAGCCGAGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 14; Length 758;  
 Best Local Similarity 99.2%; Pred. No. 1.5e-60;  
 Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 ATCAAGCGCGACCATGTGTCACATTATGCGGTTGTACAGACGATAGACCAACAGG 60  
 Db 87 ATCAAGCGCGACCATGTGTCACATTATGCGGTTGTACAGACGATAGACCAACAGG 146  
 QY 61 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGACAGAAGAGAG 120  
 Db 147 GAGTTTATGTTGAATTTGATGAAGATGAGATGTTCTATGTGATCTGACAGAAGAGAG 206  
 QY 121 ACCGTCTGCATCTGGAGAGTTTGCCAAAGCTTTTCTTTGAGCTCAGGCGGGCTG 180  
 Db 207 ACCGTCTGCATCTGGAGAGTTTGCCAAAGCTTTTCTTTGAGCTCAGGCGGGCTG 266  
 QY 181 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCTTCCAAACACACT 239  
 Db 267 GCTAACATTGCTATATTGAACAACTTGAATACCTTGATCCAGCTTCCAAACACACT 326  
 QY 240 CAGGCCACCAACCGAT 254  
 Db 327 CAGGCCACCAACCGAT 341

## RESULT 43

CB9858956  
 LOCUS CB9858956  
 DEFINITION 602568315F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:463035 5',  
 mRNA sequence.  
 ACCESSION CB9858956  
 VERSION CB9858956.1 GI:13531189  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 789)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCMI514 row: m column: 04

High quality sequence stop: 681.

FEATURES  
source

1..789  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4593035"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NIH\_MGC\_77"  
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CAGGCCATTATGGC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGGCGGCGGACATG-DT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.9  
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 789;  
Best Local Similarity 99.2%; Pred. No. 1.5e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACACAGG 60
   |||||
Db 181 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACACAGG 240
   |||||
QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACAAAGAGGAG 120
   |||||
Db 241 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGACAAAGAGGAG 300
   |||||
QY 121 ACCGCTCGCATCTGGAGAGTTGGCCAAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 180
   |||||
Db 301 ACCGCTCGCATCTGGAGAGTTGGCCAAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 360
   |||||
QY 181 GCTAACATTGCTATATTGAACAAACACTTGAATACCTTGATCCAGCTTCC-ACCACACT 239
   |||||
Db 361 GCTAACATTGCTATATTGAACAAACACTTGAATACCTTGATCCAGCTTCCACACACT 420
   |||||
QY 240 CAGGCCACCAACCGAT 254
   |||||
Db 421 CAGGCCACCAACCGAT 435
   |||||
```

RESULT 44  
BUS96174

LOCUS BUS96174 791 bp mRNA linear EST 20-SEP-2002  
DEFINITION AGENCOURT 8906879 NIH\_MGC\_142 Homo sapiens cDNA clone IMAGE:6452589  
5', mRNA sequence.

ACCESSION BUS96174.1 GI:23247933

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 791)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: NCI

cDNA Library Preparation: Michael Brownstein Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2626 row: k column: 22

High quality sequence stop: 528.

FEATURES  
source

1..791  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6452589"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NIH\_MGC\_142"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
Site 2: SfiI (ggccattatggcc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
used in cloning as follows:  
5'-AAGCATGTGTATCAAGCAGAGTGGCCATTAGCGCGGG-3' and  
5'-ATTCTAGAGCGGCGGCGGCGGACATG-DT(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the >0.5 kb  
size fraction (other fractions present in NIH\_MGC\_141).  
Library created in the laboratory of M. Brownstein (NIH,  
NIH). Note: this is a NIH\_MGC Library."

Query Match 95.0%; Score 241.4; DB 13; Length 791;  
Best Local Similarity 99.2%; Pred. No. 1.5e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACACAGG 60
   |||||
Db 136 ATCAAGCGCGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACACAGG 195
   |||||
QY 61 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGAACAAGAGGAG 120
   |||||
Db 196 GAGTTTATGTTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGAACAAGAGGAG 255
   |||||
QY 121 ACCGCTCGCATCTGGAGAGTTGGCCAAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 180
   |||||
Db 256 ACCGCTCGCATCTGGAGAGTTGGCCAAAGCCTTTTCTTTGAGGCTCAGGCGGGCTG 315
   |||||
QY 181 GCTAACATTGCTATATTGAACAAACACTTGAATACCTTGATCCAGCTTCC-ACCACACT 239
   |||||
Db 316 GCTAACATTGCTATATTGAACAAACACTTGAATACCTTGATCCAGCTTCCACACACT 375
   |||||
QY 240 CAGGCCACCAACCGAT 254
   |||||
Db 376 CAGGCCACCAACCGAT 390
   |||||
```

RESULT 45

BG757550

LOCUS

DEFINITION

BG757550 802 bp mRNA linear EST 15-MAY-2001  
602714723F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4855085 5',  
mRNA sequence.



```

ACCESSION      BI766898
VERSION        BI766898.1
KEYWORDS       GI:15758476
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          NIH-MGC http://mgs.nci.nih.gov/
JOURNAL        1 (bases 1 to 811)
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgabbs@mail.nih.gov
               Tissue Procurement: Life Technologies, Inc.
               CDNA Library Preparation: Life Technologies, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1AM1508 row: k column: 09
               High quality sequence stop: 803.

FEATURES       source
               Location/Qualifiers
               1..811
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5202944"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_122"
                /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
                anonymous pool of 24 week female lung, 16 week female
                spleen, and 20-22 week male spleens. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.4 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 026. Note:
                this is a NIH_MGC Library."

ORIGIN
Query Match      95.0%; Score 241.4; DB 12; Length 811;
Best Local Similarity 99.2%; Pred. No. 1.6e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 60
Db 126 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 185
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAGAGGAG 120
Db 186 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAGAGGAG 245
QY 121 ACCGTCGCACTCGAGAGATTGGCCAGCCCTTTTCCTTTGAGCTCAGGCGGGCTG 180
Db 246 ACCGTCGCACTCGAGAGATTGGCCAGCCCTTTTCCTTTGAGCTCAGGCGGGCTG 305
QY 181 GCTAACATGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239
Db 306 GCTAACATGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 365
QY 240 CAGGCCACCAACCGAT 254
Db 366 CAGGCCACCAACCGAT 380

RESULT 48
BI668794
LOCUS          821 bp mRNA linear EST 12-SEP-2001
DEFINITION     603294742F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5313754 5',
               mRNA sequence.
ACCESSION      BI668794

BI766898
VERSION        BI766898.1
KEYWORDS       GI:15758476
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          NIH-MGC http://mgs.nci.nih.gov/
JOURNAL        1 (bases 1 to 821)
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgabbs@mail.nih.gov
               Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
               CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
               Toshiyuki and Piero Carninci (RIKEN)
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1AM1795 row: d column: 11
               High quality sequence stop: 805.

FEATURES       source
               Location/Qualifiers
               1..821
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5313754"
                /tissue_type="hypothalamus"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_96"
                /note="Organ: brain; Vector: pBluescriptR (modified
                pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
                (gtcag); Oligo-dT primed using primer
                5'-TTTTTTTTTTTTTNN-3', size-selected for average
                insert size 2.3 kb and normalized to 500 ng. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NHGRI/NHGRI, National Institutes of Health). Note: this is
                a NIH_MGC Library."

ORIGIN
Query Match      95.0%; Score 241.4; DB 12; Length 821;
Best Local Similarity 99.2%; Pred. No. 1.6e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 60
Db 153 ATCAAGCGGACCATGTGTCACACTTATGCGCGTTTGTACAGACGATAGACCAACAGGG 212
QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAGAGGAG 120
Db 213 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAGAGGAG 272
QY 121 ACCGTCGCACTCGAGAGATTGGCCAGCCCTTTTCCTTTGAGCTCAGGCGGGCTG 180
Db 273 ACCGTCGCACTCGAGAGATTGGCCAGCCCTTTTCCTTTGAGCTCAGGCGGGCTG 332
QY 181 GCTAACATGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 239
Db 333 GCTAACATGCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACAT 392
QY 240 CAGGCCACCAACCGAT 254
Db 393 CAGGCCACCAACCGAT 407

RESULT 49
BI668794
LOCUS          826 bp mRNA linear EST 14-MAR-2001
DEFINITION     602498354F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:461196 5',
               mRNA sequence.

```

```

ACCESSION      BG431923
VERSION        BG431923.1  GI:13338429
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 826)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
AUTHORS        Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: CLONTECH Laboratories, Inc.
              CDNA Library Preparation: CLONTECH Laboratories, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM1359 row: 1 column: 13
              High quality sequence stop: 646.
FEATURES       source
              1..826
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4611996"
              /lab_host="DH10B (T1 phage-resistant)"
              /notes="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site 1:
              SfII (ggcgctcgccgc); Site 2: SfII (ggcattatggcc); 5' and
              3' adaptors were used in cloning as follows: 5' adaptor
              sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
              5'-ATTCTAGAGCCGAGGCGGCATG-dt(30)BN-3 (where B = A,
              C, or G and N = A, C, G, or T). Average insert size 1.65
              kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
              by PCR. This library was enriched for full-length clones
              and was constructed by Clontech Laboratories (Palo Alto,
              CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match      95.0%; Score 241.4; DB 12; Length 826;
Best Local Similarity 99.2%; Pred. No. 1.6e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1 ATCAAGCGGACCATGTGTCAACTTATGCCCGCTTTGTACAGCGCATAGACCAACAGGG 60
        |||||
Db      174 ATCAAGCGGACCATGTGTCAACTTATGCCCGCTTTGTACAGCGCATAGACCAACAGGG 233

QY      61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAAGGAG 120
        |||||
Db      234 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAAGGAG 293

QY      121 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAGACCTTTTCCTTTGAGGCTCAGGCGGGCTG 180
        |||||
Db      294 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAGACCTTTTCCTTTGAGGCTCAGGCGGGCTG 353

QY      181 GCTAACATTCCTATTATTTGAACAACAATTGAATACCTTTGATCCAGCGTTCC-ACCACACT 239
        |||||
Db      354 GCTAACATTCCTATTATTTGAACAACAATTGAATACCTTTGATCCAGCGTTCCACACACT 413

QY      240 CAGGCCACCAACCGAT 254
        |||||
Db      414 CAGGCCACCAACCGAT 428

RESULT 50
CB994002
LOCUS
DEFINITION AGNCOURT_136173:0 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30333979 5', mRNA sequence.
ACCESSION CB994002

```

```

CB994002.1  GI:30289522
EST.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 830)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
AUTHORS      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: Dr. Stefan Hansson
              CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
              and advice from Piero Carninci (RIKEN)
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: NDAM353 row: h column: 20
              High quality sequence stop: 577.
FEATURES     source
              1..830
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:30333979"
              /tissue_type="pre-eclamptic placenta"
              /lab_host="DH10B Tona"
              /clone_lib="NIH_MGC_148"
              /note="Organ: placenta; Vector: pBluescriptR; Site 1:
              all-XhoI; Site 2: BamH; Library is oligo-dT primed and
              directionally cloned using primer
              5'-TTTTTTTTTTTTTNN-3', size-selected for average insert
              size 2.3 kb and normalized to 80T 5. This is a primary
              library enriched for full-length clones and constructed
              using the Cap-trapper method (Carninci, in preparation).
              Library constructed by M. Brownstein (NHGRI/NIH),
              National Institutes of Health). Note: this is a NIH_MGC
              Library."
ORIGIN
Query Match      95.0%; Score 241.4; DB 14; Length 830;
Best Local Similarity 99.2%; Pred. No. 1.6e-60;
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1 ATCAAGCGGACCATGTGTCAACTTATGCCCGCTTTGTACAGCGCATAGACCAACAGGG 60
        |||||
Db      174 ATCAAGCGGACCATGTGTCAACTTATGCCCGCTTTGTACAGCGCATAGACCAACAGGG 233

QY      61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAAGGAG 120
        |||||
Db      234 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGATCTGGACAGAAGGAG 293

QY      121 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAGACCTTTTCCTTTGAGGCTCAGGCGGGCTG 180
        |||||
Db      294 ACCGTCCTGGCATCTGGAGGAGTTTGGCCAGACCTTTTCCTTTGAGGCTCAGGCGGGCTG 353

QY      181 GCTAACATTCCTATTATTTGAACAACAATTGAATACCTTTGATCCAGCGTTCC-ACCACACT 239
        |||||
Db      354 GCTAACATTCCTATTATTTGAACAACAATTGAATACCTTTGATCCAGCGTTCCACACACT 413

QY      240 CAGGCCACCAACCGAT 254
        |||||
Db      414 CAGGCCACCAACCGAT 428

RESULT 51
CB958376
LOCUS
DEFINITION AGNCOURT_136667:61 NIH_MGC_184 Homo sapiens cDNA clone
IMAGE:30353204 5', mRNA sequence.

```

```

DEFINITION      603042205FL NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182560 5'
rna sequence.
ACCESSION      B1517599
VERSION        B1517599.1  GI:15342391
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
                1 (bases 1 to 837)
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999).
AUTHORS        Contact: Robert Strausberg, Ph.D.
                Email: cgabbs@mail.nih.gov
JOURNAL        Tissue Procurement: Life Technologies, Inc.
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
COMMENT        Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM11455 row: j column: 01
                High quality sequence stop: 833.

```

```

FEATURES
source
      1. .837
      Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5182560"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_116"
        /notes="Organ: pooled colon, kidney, stomach; Vector:
        pCMV-SPORT6; Site:1: NotI; Site2: EcoRV (destroyed); RNA
        source anonymous pool of 3 colons, age 26 yo male, 49 yo
        female, 71 yo male colon; 46 yo male kidney, and pool of 2
        stomachs, 62 yo male and 70 yo female. Library is
        oligo-dT primed and directionally cloned (EcoRV site is
        destroyed upon cloning). Average insert size 1.4 kb,
        insert size range 1-3 kb. Library is normalized and
        enriched for full-length clones and was constructed by C.
        Gruber (Invitrogen). Research Genetics tracking code
        023. Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	95.0%	Score 241.4;	DB 12;	Length 837;
Best Local Similarity	99.2%	Pred. No. 1.6e-60;		
Mismatches	253;	Conservative	1;	Indels 1;
Gaps	1;			

Qy	1	ATCAAGCGGACCATGTGTCAACTTATGCCCGTTGTACAGACGCATAGACCCACACGG	60
Db	152	ATCAAGCGGACCATGTGTCAACTTATGCCCGTTGTACAGACGCATAGACCAACACGG	211
Qy	61	GAGTTTATGTTTGAATTGTAGTAAGATGAGATGTTCTATGTGGATCTGGACAAGAAGAG	120
Db	212	GAGTTTATGTTTGAATTGTAGTAAGATGAGATGTTCTATGTGGATCTGGACAAGAAGAG	271
Qy	121	ACCGTCTGGGACTCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG	180
Db	272	ACCGTCTGGCCTCTGGAGGAGTTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGGCTG	331
Qy	181	GCTAACATTTGCTATATTGAAACAACAACCTTGAATACCTTGTATCCAGCGTTCC	239
Db	332	GCTAACATTTGCTATATTGAAACAACAACCTTGAATACCTTGTATCCAGCGTTCCACACACT	391
Qy	240	CAGGCCACCAACCGAT	254
Db	392	CAGGCCACCAACGAT	406

RESULT 53			
BI261592			
LOCUS			
RT261592	843 bp	mRNA	linear
			EST 17-JUL-2001

DEFINITION 602953662F1 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5087774 5', mRNA sequence.

ACCESSION B1261592

VERSION B1261592.1 GI:14820975

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LCM1845 row: 1 column: 15  
High quality sequence stop: 833.

## FEATURES

Location/Qualifiers  
1..843  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5087774"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 12; Length 843;  
Best Local Similarity 99.2%; Pred. No. 1.6e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60  
DB 134 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 133

QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 120  
DB 194 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 253

QY 121 ACCGTCGGCATCTGGAGGAGTTGGCCAGCGCTTTCCTTTGAGGCTCAGGCGGGCTG 180  
DB 254 ACCGTCGGCATCTGGAGGAGTTGGCCAGCGCTTTCCTTTGAGGCTCAGGCGGGCTG 313

QY 181 GCTAACATTCGTATTATTAACCAACTTGAATACCTTGTATCCAGCGTTCC-ACCACACT 239  
DB 314 GCTAACATTCGTATTATTAACCAACTTGAATACCTTGTATCCAGCGTTCCACCACT 373

QY 240 CAGGCCACCAACCGAT 254  
DB 374 CAGGCCACCAACCGAT 388

## RESULT 54

CD558780  
LOCUS EST 11-JUN-2003  
DEFINITION AGENCOURT\_14400797 NIH\_MGC\_181 Homo sapiens cDNA clone

IMAGE:30396983 5', mRNA sequence.

CD558780

CD558780.1 GI:31584848

EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: ND4479 row: 1 column: 24  
High quality sequence stop: 671.

## FEATURES

Location/Qualifiers  
1..851  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30396983"  
/tissue\_type="White Matter"  
/dev\_stage="Unknown"  
/lab\_host="DH10B-Ton A (T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_181"  
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by Invitrogen). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 241.4; DB 14; Length 851;  
Best Local Similarity 99.2%; Pred. No. 1.6e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 60  
DB 144 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGCGCATAGACCAACAGGG 203

QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 120  
DB 204 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTCGATCTGGACAAGAGGAG 263

QY 121 ACCGTCGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 180  
DB 264 ACCGTCGGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGCGGGCTG 323

QY 181 GCTAACATTCGTATTATTAACCAACTTGAATACCTTGTATCCAGCGTTCC-ACCACACT 239  
DB 324 GCTAACATTCGTATTATTAACCAACTTGAATACCTTGTATCCAGCGTTCCACCACT 383

QY 240 CAGGCCACCAACCGAT 254  
DB 384 CAGGCCACCAACCGAT 398

## RESULT 55

BQ890384  
LOCUS EST 16-AUG-2002  
DEFINITION AGENCOURT\_8585789 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:6304604  
5', mRNA sequence.

ACCESSION BQ890384  
VERSION BQ890384.1 GI:22282398  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 857)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM2525 row: i column: 21  
High quality sequence stop: 617.

FEATURES  
source  
1..857  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6304604"  
/tissue type="lymphoma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 95"  
/note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

ORIGIN  
Query Match 95.0%; Score 241.4; DB 13; Length 857;  
Best Local Similarity 99.2%; Pred. No. 1.6e-60;  
Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGGATAGACCAACAGGG 60  
DB 145 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGGATAGACCAACAGGG 204  
QY 61 GAGTTTATGTTGAATTGATGAAGATGATGTTCTATGCTGATCTGGACAGAGGAG 120  
DB 205 GAGTTTATGTTGAATTGATGAAGATGATGTTCTATGCTGATCTGGACAGAGGAG 264  
QY 121 ACCGCTCGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGCTG 180  
DB 265 ACCGCTCGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGCTG 324  
QY 181 GCTAACATTTGCTATTTGAAACAACTTGAATACCTTGATCCAGGTTCC-ACCACAT 239  
DB 325 GCTAACATTTGCTATTTGAAACAACTTGAATACCTTGATCCAGGTTCCACACCAT 384  
QY 240 CAGGCCACCAACCAT 254  
DB 385 CAGGCCACCAACCAT 399

RESULT 56  
BQ721555  
LOCUS BQ721555 871 bp mRNA linear EST 25-SEP-2001  
DEFINITION IMAGEN:6187825 5', mRNA sequence.  
ACCESSION BQ721555

VERSION BI767351.1 GI:15758929  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 871)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM11519 row: e column: 22  
High quality sequence stop: 850.

FEATURES  
source  
1..871  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5207037"  
/lab host="DH10B"  
/clone lib="NIH MGC 122"  
/note="Organ: Pooled lung and spleen; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH\_MGC Library."

ORIGIN  
Query Match 95.0%; Score 241.4; DB 12; Length 871;  
Best Local Similarity 99.2%; Pred. No. 1.6e-60;  
Matches 233; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGGATAGACCAACAGGG 60  
DB 134 ATCAAGCGGACCATGTGTCAACTTATGCGCGTTGTACAGACGGATAGACCAACAGGG 193  
QY 61 GAGTTTATGTTGAATTGATGAAGATGATGTTCTATGCTGATCTGGACAGAGGAG 120  
DB 194 GAGTTTATGTTGAATTGATGAAGATGATGTTCTATGCTGATCTGGACAGAGGAG 253  
QY 121 ACCGCTCGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGCTG 180  
DB 254 ACCGCTCGCATCTGGAGGAGTTGGCCAAAGCCTTTTCCTTTGAGGCTCAGGGCGGCTG 313  
QY 181 GCTAACATTTGCTATTTGAAACAACTTGAATACCTTGATCCAGGTTCC-ACCACAT 239  
DB 314 GCTAACATTTGCTATTTGAAACAACTTGAATACCTTGATCCAGGTTCCACACCAT 373  
QY 240 CAGGCCACCAACCAT 254  
DB 374 CAGGCCACCAACCAT 388

RESULT 57  
BQ721555  
LOCUS BQ721555 872 bp mRNA linear EST 16-JUL-2002  
DEFINITION IMAGEN:6187825 5', mRNA sequence.  
ACCESSION BQ721555  
VERSION BQ721555.1 GI:21860452



KEYWORDS SOURCE ORGANISM	EST. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 872)
REFERENCE AUTHORS TITLE JOURNAL COMMENT	NIH-MGC <a href="http://mgs.nci.nih.gov/">http://mgs.nci.nih.gov/</a> National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapsb@mail.nih.gov">cgapsb@mail.nih.gov</a> Tissue procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L1M13582 row: h column: 02 High quality sequence stop: 678. Location/Qualifiers 1. .872 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6187825" /sex="male" /tissue_type="sympathetic trunk" /dev_stage="adult, 16 yr" /lab_host="DH10B" /clone_lib="Lupski sympathetic trunk" /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCATCGCTCCG-3' and 5'-GACTAGTCTAGATCGGCGCCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."
FEATURES source	1. .872 Location/Qualifiers /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6187825" /sex="male" /tissue_type="sympathetic trunk" /dev_stage="adult, 16 yr" /lab_host="DH10B" /clone_lib="Lupski sympathetic trunk" /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCATCGCTCCG-3' and 5'-GACTAGTCTAGATCGGCGCCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."
ORIGIN	Query Match 95.0%; Score 241.4; DB 13; Length 872; Best Local Similarity 99.2%; Pred. No. 1.6e-60; Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1; QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGG 60 DB 147 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGG 206 QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGGAG 120 DB 207 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGGAG 266 QY 121 ACCGTCGGCATCTGGAGGAGTTGGCCAGCCCTTTTCCTTTGAGGCTCAGGGCGGCTG 180 DB 267 ACCGTCGGCATCTGGAGGAGTTGGCCAGCCCTTTTCCTTTGAGGCTCAGGGCGGCTG 326 QY 181 GCTAACATTCCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239 DB 327 GCTAACATTCCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 386 QY 240 CAGGCCACCAACCGAT 254 DB 387 CAGGCCACCAACCGAT 401
RESULT 58	CG540219
LOCUS	BG540219
DEFINITION	602569108F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4693849 5', mRNA sequence.

ACCESSION	BG540219
VERSION	BG540219.1 GI:13532452
KEYWORDS	EST.
SOURCE	Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 873)
REFERENCE	NIH-MGC <a href="http://mgs.nci.nih.gov/">http://mgs.nci.nih.gov/</a> National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapsb@mail.nih.gov">cgapsb@mail.nih.gov</a> Tissue procurement: CLONTECH Laboratories, Inc. CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLCM1516 row: o column: 02 High quality sequence stop: 752. Location/Qualifiers 1. .873 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4693849" /lab_host="DH103 (T1 phage-resistant)" /clone_lib="NIH_MGC_77" /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcgattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCGCATATGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGGCGCGCGCACATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
ORIGIN	Query Match 95.0%; Score 241.4; DB 12; Length 873; Best Local Similarity 99.2%; Pred. No. 1.6e-60; Matches 253; Conservative 0; Mismatches 1; Indels 1; Gaps 1; QY 1 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGG 60 DB 146 ATCAAGCGGACCATGTGTCAACTTATGCGCGCTTTGTACAGACGCATAGACCAACAGGG 205 QY 61 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGGAG 120 DB 206 GAGTTTATGTTGAATTGATGAAGATGAGATGTTCTATGTGGATCTGGAACAAGAGGAG 265 QY 121 ACCGTCGGCATCTGGAGGAGTTGGCCAGCCCTTTTCCTTTGAGGCTCAGGGCGGCTG 180 DB 266 ACCGTCGGCATCTGGAGGAGTTGGCCAGCCCTTTTCCTTTGAGGCTCAGGGCGGCTG 325 QY 181 GCTAACATTCCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 239 DB 326 GCTAACATTCCTATATTGAACAACTTGAATACCTTGATCCAGCGTTCC-ACCACACT 385 QY 240 CAGGCCACCAACCGAT 254 DB 386 CAGGCCACCAACCGAT 400
RESULT 59	CB986657
LOCUS	CB986657
DEFINITION	AGENCOURT 13567736 NIH_MGC_184 Homo sapiens cDNA clone IMAGE:30326868 5', mRNA sequence.
ACCESSION	CB986657



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Computer Ltd.

OM nucleic - nucleic search, using sw model

Rur on: April 20, 2004, 02:07:57 : Search time 128.353 Seconds

(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819B-38

Perfect score: 20

Sequence: 1 ggcttggagcgctctttta 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373663 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : N Geneseq\_29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	20	100.0	20	6	AAL48182 Human HLA
C 2	20	100.0	20	6	AAL48202 Human HLA
C 3	20	100.0	20	6	AAL48200 Human HLA
C 4	20	100.0	20	6	AAL48166 Human HLA
C 5	18	90.0	545	5	AS64515 DNA encod
C 6	18	90.0	2076	5	AS64519 DNA encod
C 7	16.4	82.0	349980	5	AHL41224 Pyrococcus
C 8	16	80.0	258	3	AAC18777 Human sec
C 9	16	80.0	1800	2	AAV27582 Helicobac
C 10	15.8	79.0	582	8	ADA49275 Maize gen
C 11	15.8	79.0	648	6	ABX66962 Helicobac
C 12	15.8	79.0	673	6	ABX66528 Helicobac
C 13	15.8	79.0	759	7	ADA71175 Rice gene
C 14	15.8	79.0	1024	6	ABX66344 Helicobac
C 15	15.8	79.0	1024	6	ABX65416 Helicobac
C 16	15.8	79.0	1061	4	ABL02097 Drosophil
C 17	15.8	79.0	1110	4	AS53627 Helicobac
C 18	15.8	79.0	1110	7	ACH34729 Prokaryot
C 19	15.8	79.0	1170	2	AAX14331 H. pylori
C 20	15.8	79.0	1581	7	ACA51800 Prokaryot
C 21	15.8	79.0	3581	4	ABL02096 Drosophil
C 22	15.8	79.0	8546	2	AAX13083 Enterococ
C 23	15.8	79.0	8546	6	ABS98878 Enterococ

C 24	15.8	79.0	349980	5	AAH68528	Aah68528 C glutami
C 25	15.4	77.0	224	5	ABV19564	Abv19564 Human pro
C 26	15.4	77.0	236	5	ABV49725	Abv49725 Human pro
C 27	15.4	77.0	1176	6	AAQ39110	Aaq39110 Human lun
C 28	15.4	77.0	1203	4	ABL26151	AbL26151 Drosophil
C 29	15.4	77.0	1440	2	AAAX14179	Aax14179 H. pylori
C 30	15.4	77.0	3402	4	ABL10673	AbL10673 Drosophil
C 31	15.4	77.0	5583	4	ABL10672	AbL10672 Drosophil
C 32	15.4	77.0	10766	4	AAK90312	Aak90312 Human dig
C 33	15.4	77.0	10766	5	AAK98955	Aak98955 Genomic s
C 34	15.4	77.0	10766	8	ADB32855	Adb32855 Human nov
C 35	15.2	76.0	46	2	AAAX26508	Aax26508 WO 990919
C 36	15.2	76.0	187	6	ABL79933	AbL79933 Human ova
C 37	15.2	76.0	414	8	ACH18032	Ach18032 Human adu
C 38	15.2	76.0	422	5	AAAG6006	Aag6006 DNA encod
C 39	15.2	76.0	627	2	AAQ87825	Aaq87825 Agmenellu
C 40	15.2	76.0	747	5	AAH67991	Aah67991 C glutami
C 41	15.2	76.0	762	6	ABX74665	Abx74665 Bacillus
C 42	15.2	76.0	765	2	AAV24610	Aav24610 H. pylori
C 43	15.2	76.0	768	2	AAV24886	Aav24886 H. pylori
C 44	15.2	76.0	768	2	AAH75824	Aah75824 H. pylori
C 45	15.2	76.0	1110	4	AAAS3795	Aas3795 Helicobac
C 46	15.2	76.0	1209	7	ADA70113	Ada70113 Rice gene
C 47	15.2	76.0	1686	8	ADB09055	Adb09055 Allostococ
C 48	15.2	76.0	1686	8	ADB09057	Adb09057 Allostococ
C 49	15.2	76.0	1686	8	ADB09059	Adb09059 Allostococ
C 50	15.2	76.0	1727	2	AAQ11036	Aaq11036 Gene encod
C 51	15.2	76.0	1838	7	ACD19454	Acd19454 cDNA encod
C 52	15.2	76.0	2044	1	AAN81505	Aan81505 DNA encod
C 53	15.2	76.0	2061	4	AAAS3803	Aas3803 Helicobac
C 54	15.2	76.0	2136	2	AAQ26642	Aaq26642 ISRml in
C 55	15.2	76.0	2253	5	AAH67990	Aah67990 C glutami
C 56	15.2	76.0	2466	7	ACA00124	Aca00124 C. glutam
C 57	15.2	76.0	2586	9	ADD13666	Add13666 C. glutam
C 58	15.2	76.0	2884	4	AHL4544	Ahl4544 Human cDN
C 59	15.2	76.0	3159	9	ADD45409	Add45409 Rat gene
C 60	15.2	76.0	3240	4	ABL02064	AbL02064 Drosophil
C 61	15.2	76.0	3241	4	ABL20820	AbL20820 Drosophil
C 62	15.2	76.0	3241	4	ABL20822	AbL20822 Drosophil
C 63	15.2	76.0	3411	2	AAAT28950	Aat28950 Helicobac
C 64	15.2	76.0	4018	7	ABZ34827	Abz34827 Coding se
C 65	15.2	76.0	4234	2	AAQ87819	Aaq87819 Agmenellu
C 66	15.2	76.0	8396	3	AAZ59353	Aaz59353 Human STP
C 67	15.2	76.0	8672	9	ADD18777	Add18777 Human dis
C 68	15.2	76.0	34503	4	AAK79621	Aak79621 Human imm
C 69	15.2	76.0	34503	4	AAK57032	Aak57032 Human imm
C 70	15.2	76.0	35100	2	AAV73803	Aav73803 KSHV LTR
C 71	15.2	76.0	48551	6	AAAS20800	Aas20800 Clostridi
C 72	15.2	76.0	110000	8	ADB12064_07	Continuation (8 of
C 73	15.2	76.0	137507	2	AAV19941	Aav19941 KSHV long
C 74	15.2	76.0	349980	5	AAH68533	Aah68533 C glutami
C 75	15.2	76.0	349980	5	AAAF96431	Aaf96431 Pyrococci
C 76	15	75.0	66	5	AAAD18666	Aad18666 Porcine f
C 77	15	75.0	105	5	AAAD18663	Aad18663 Modified
C 78	15	75.0	2266	4	ABL24228	AbL24228 Drosophil
C 79	15	75.0	3066	4	ABL11438	AbL11438 Drosophil
C 80	15	75.0	3152	4	ABL13448	AbL13448 Drosophil
C 81	15	75.0	4404	5	ABK18667	Abk18667 Modified
C 82	14.8	74.0	301	6	ABK80949	Abk80949 Bacillus
C 83	14.8	74.0	352	4	AAI15213	Aai15213 Probe #51
C 84	14.8	74.0	352	4	AAK05059	Aak05059 Human bra
C 85	14.8	74.0	367	5	ABS05284	Abs05284 Human gen
C 86	14.8	74.0	367	5	ABV56924	Abv56924 Human pro
C 87	14.8	74.0	386	4	AAI26325	Aai26325 Human foe
C 88	14.8	74.0	386	4	ABA74125	Aba74125 Human probe #23
C 89	14.8	74.0	386	4	AAI54577	Aai54577 Probe #17
C 90	14.8	74.0	386	4	ABA39142	AbA39142 Probe #17
C 91	14.8	74.0	386	4	AAK48747	Aak48747 Human bon
C 92	14.8	74.0	386	4	AAK22578	Aak22578 Human bra
C 93	14.8	74.0	386	4	ABS48421	Abs48421 Human liv
C 94	14.8	74.0	386	6	ABS22461	Abs22461 Human gen
C 95	14.8	74.0	408	4	AAI24392	Aai24392 Probe #14
C 96	14.8	74.0	408	4	AAK17778	Aak17778 Human bra

97 14.8 74.0 408 6 ABS17785 Abs17785 Human gen  
 98 14.8 74.0 480 5 ABA06274 ABA06274 Soy bean  
 99 14.8 74.0 480 5 ABA06273 ABA06273 Soy bean  
 100 14.8 74.0 481 9 ADB51204 ADB51204 Primary r  
 101 14.8 74.0 534 4 AAI17148 AAI17148 Probe #70  
 102 14.8 74.0 534 4 ABA61630 ABA61630 Human foe  
 103 14.8 74.0 534 4 AAI41542 AAI41542 Probe #10  
 104 14.8 74.0 534 4 ABA29292 ABA29292 Probe #77  
 105 14.8 74.0 534 4 AAK35825 AAK35825 Human bon  
 106 14.8 74.0 534 4 AAK09930 AAK09930 Human bra  
 107 14.8 74.0 534 4 ABS35535 ABS35535 Human liv  
 108 14.8 74.0 534 6 ABS10060 ABS10060 Human gen  
 109 14.8 74.0 550 9 ADD34331 ADD34331 Mouse mit  
 110 14.8 74.0 637 9 ADD34627 ADD34627 Mouse mit  
 111 14.8 74.0 765 5 AAB65040 AAB65040 C glutami  
 112 14.8 74.0 777 6 ABO67928 ABO67928 Listeria  
 113 14.8 74.0 868 3 AAC47508 AAC47508 Arabidops  
 114 14.8 74.0 870 3 AAC34139 AAC34139 Arabidops  
 115 14.8 74.0 980 5 AAS83196 AAS83196 DNA enco  
 116 14.8 74.0 1050 4 AAI17172 AAI17172 ActA prot  
 117 14.8 74.0 1128 4 AAI17171 AAI17171 ActA prot  
 118 14.8 74.0 1156 6 ABK63440 ABK63440 Rat seque  
 119 14.8 74.0 1156 9 ADB57852 ADB57852 Toxicity-  
 120 14.8 74.0 1156 9 ADB52358 ADB52358 Primary r  
 121 14.8 74.0 1156 9 ADB85151 ADB85151 Rat UDP-g  
 122 14.8 74.0 1229 4 ABL05839 ABL05839 Drosophil  
 123 14.8 74.0 1467 5 AAB65041 AAB65041 C glutami  
 124 14.8 74.0 1830 4 AAI17170 AAI17170 ActA prot  
 125 14.8 74.0 1836 7 ABT42297 ABT42297 Toxicity  
 126 14.8 74.0 1841 4 AAB14772 AAB14772 Human cDN  
 127 14.8 74.0 1878 5 AAS73554 AAS73554 DNA enco  
 128 14.8 74.0 1920 2 AAV37026 AAV37026 Listeria  
 129 14.8 74.0 1920 6 ABO67930 ABO67930 Listeria  
 130 14.8 74.0 1920 6 ABO67930 ABO67930 Listeria  
 131 14.8 74.0 2016 5 AAS73555 AAS73555 DNA enco  
 132 14.8 74.0 2631 7 ACA53272 ACA53272 Prokaryot  
 133 14.8 74.0 2726 4 AAS26593 AAS26593 Human cDN  
 134 14.8 74.0 2726 7 ABX73934 ABX73934 Human ro  
 135 14.8 74.0 2744 4 AAS26178 AAS26178 Human cDN  
 136 14.8 74.0 2744 7 ABX73519 ABX73519 Human ro  
 137 14.8 74.0 2850 2 AAV64278 AAV64278 Human EGR  
 138 14.8 74.0 2911 6 ABX82508 ABX82508 Human sec  
 139 14.8 74.0 3074 7 ACA31588 ACA31588 Prokaryot  
 140 14.8 74.0 3293 4 ABL05838 ABL05838 Drosophil  
 141 14.8 74.0 3444 7 ACS35695 ACS35695 Prokaryot  
 142 14.8 74.0 3587 4 ABL23868 ABL23868 Drosophil  
 143 14.8 74.0 3587 7 ACA35793 ACA35793 Prokaryot  
 144 14.8 74.0 4272 6 ABX11094 ABX11094 Human bre  
 145 14.8 74.0 4272 7 ABX10965 ABX10965 cDNA enco  
 146 14.8 74.0 4338 9 ADE25700 ADE25700 Human cDN  
 147 14.8 74.0 4339 7 ABX62934 ABX62934 Human act  
 148 14.8 74.0 4724 2 AAX20542 AAX20542 Polynucle  
 149 14.8 74.0 6714 4 ABL16321 ABL16321 Drosophil  
 150 14.8 74.0 7272 9 ADB56211 ADB56211 Toxicity-  
 151 14.8 74.0 7471 3 AAZ88789 AAZ88789 B. subtil  
 152 14.8 74.0 11838 4 ABL16320 ABL16320 Drosophil  
 153 14.8 74.0 110000 6 ABA03041\_02 ABA03041\_02  
 154 14.8 74.0 145831 6 ABL69213 ABL69213 Prostate  
 155 14.8 74.0 145831 6 ABL66806 ABL66806 Lung canc  
 156 14.8 74.0 145831 6 ABL68588 ABL68588 Kidney ca  
 157 14.8 74.0 145831 6 ABL62309 ABL62309 Colon ade  
 158 14.8 74.0 145831 6 ABL10149 ABL10149 Human bre  
 159 14.8 74.0 349380 5 AAB64966 AAB64966 C glutami  
 160 14.4 72.0 126 4 ABA69371 ABA69371 Human foe  
 161 14.4 72.0 126 4 ABA36307 ABA36307 Probe #14  
 162 14.4 72.0 126 4 ABA36306 ABA36306 Human liv  
 163 14.4 72.0 126 6 ABS17584 ABS17584 Human gen  
 164 14.4 72.0 186 7 ACA23632 ACA23632 Prokaryot  
 165 14.4 72.0 195 7 ACA26374 ACA26374 Prokaryot  
 166 14.4 72.0 213 4 AAI24080 AAI24080 Probe #14  
 167 14.4 72.0 213 4 ABA69203 ABA69203 Human foe  
 168 14.4 72.0 213 4 AAI49375 AAI49375 Probe #18  
 169 14.4 72.0 213 4 ABA51196 ABA51196 Human bre

## ALIGNMENTS

## RESULT 1

AAL48182/c

ID AAL48182 standard; DNA; 20 BP.

XX AC AAL48182;

XX DT 01-OCT-2002 (first entry)

XX DE Human HLA DPB1 locus polymorphism multiplex capture sequence #2.

XX KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;

XX KW flow cytometry; human; DPB1; capture tag; ss.

XX OS Homo sapiens.

XX PN WO200194639-A1.

XX PD 13-DEC-2001.

XX PF 07-JUN-2001; 2001WO-US018590.

XX PR 08-JUN-2000; 2000US-0210759P.

XX XX (REGC ) UNIV CALIFORNIA.

XX PI White PS, Torney DC;

XX WPI; 2002-566450/60.

XX PT Identifying sequences useful as address/capture tags for flow cytometry

XX PT based minisequencing, by generating tag sequences and rejecting sequences

XX XX based on certain parameters e.g. sequences which form stable hairpins.

XX PS Disclosure; Page 9; 35pp; English.

XX CC The present invention relates to a method of identifying sequences useful

XX CC as address/capture tags, involving rejecting sequences having common sub-

XX CC sequences with a sub-sequence length greater than specified number of

XX CC bases, and sequences which can form stable hairpins and stable dimers

XX CC from a sample of oligonucleotides, and selecting those sequences in the

XX CC sample that would hybridize to their respective complements with a high

XX CC degree of specificity. The method is useful for identifying a set of

XX CC sequences useful as address/capture tags which can be used for

XX CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow

XX CC cytometry assay. The present sequence is a capture tag described in the

XX CC exemplification of the invention

XX SQ Sequence 20 BP; 7 A; 6 C; 4 G; 3 T; 0 U; 0 Other;

XX

Query Match 100.0%; Score 20; DB 6; Length 20;

Best Local Similarity 100.0%; Pred.No. 0.92;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTAA 20

DB 20 GCCTTTGGAGCGCTCTTAA 1

```

RESULT 2
AAL48202
ID AAL48202 standard; DNA; 20 BP.
XX AC AAL48202;
XX DT 01-OCT-2002 (first entry)
XX DE Human HLA DPB1 locus polymorphism address tag sequence #2.
XX KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX KW flow cytometry; human; DPB1; address tag; ss.
XX OS Homo sapiens.
XX PN WO200194639-A1.
XX PD 13-DEC-2001.
XX PF 07-JUN-2001; 2001WO-US018590.
XX PR 08-JUN-2000; 2000US-0210759P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI White PS, Torney DC;
XX DR WPI; 2002-566450/60.
XX PT Identifying sequences useful as address/capture tags for flow cytometry
XX PT based minisequencing, by generating tag sequences and rejecting sequences
XX PT based on certain parameters e.g. sequences which form stable hairpins.
XX PS Disclosure; Page 14; 35pp; English.
XX CC The present invention relates to a method of identifying sequences useful
XX CC as address/capture tags, involving rejecting sequences having common sub-
XX CC sequences with a sub-sequence length greater than specified number of
XX CC bases, and sequences which can form stable hairpins and stable dimers
XX CC from a sample of oligonucleotides, and selecting those sequences in the
XX CC sample that would hybridise to their respective complements with a high
XX CC degree of specificity. The method is useful for identifying a set of
XX CC sequences useful as address/capture tags which can be used for
XX CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
XX CC cytometry assay. The present sequence is an address tag described in the
XX CC exemplification of the invention
XX SQ Sequence 20 BP; 3 A; 4 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTGAGCGCTCTTTAA 20
Db 1 GGCCTTGAGCGCTCTTTAA 20

RESULT 3
AAL48200/c
ID AAL48200 standard; DNA; 20 BP.
XX AC AAL48200;
XX DT 01-OCT-2002 (first entry)
XX DE Human HLA DPB1 locus polymorphism multiplex capture sequence #2.
XX KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX KW flow cytometry; human; DPB1; capture tag; ss.
XX OS Homo sapiens.

```

```

XX WO200194639-A1.
XX 13-DEC-2001.
XX 07-JUN-2001; 2001WO-US018590.
XX 08-JUN-2000; 2000US-0210759P.
XX (REGC ) UNIV CALIFORNIA.
XX White PS, Torney DC;
XX WPI; 2002-566450/60.
XX Identifying sequences useful as address/capture tags for flow cytometry
XX based minisequencing, by generating tag sequences and rejecting sequences
XX based on certain parameters e.g. sequences which form stable hairpins.
XX Disclosure; Page 14; 35pp; English.
XX CC The present invention relates to a method of identifying sequences useful
XX CC as address/capture tags, involving rejecting sequences having common sub-
XX CC sequences with a sub-sequence length greater than specified number of
XX CC bases, and sequences which can form stable hairpins and stable dimers
XX CC from a sample of oligonucleotides, and selecting those sequences in the
XX CC sample that would hybridise to their respective complements with a high
XX CC degree of specificity. The method is useful for identifying a set of
XX CC sequences useful as address/capture tags which can be used for
XX CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
XX CC cytometry assay. The present sequence is a capture tag described in the
XX CC exemplification of the invention
XX SQ Sequence 20 BP; 7 A; 6 C; 4 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTGAGCGCTCTTTAA 20
Db 20 GGCCTTGAGCGCTCTTTAA 1

RESULT 4
AAL48166
ID AAL48166 standard; DNA; 20 BP.
XX AC AAL48166;
XX DT 01-OCT-2002 (first entry)
XX DE Human HLA DPB1 locus polymorphism address tag sequence #2.
XX KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX KW flow cytometry; human; DPB1; capture tag; ss.
XX OS Homo sapiens.
XX PN WO200194639-A1.
XX PD 13-DEC-2001.
XX PF 07-JUN-2001; 2001WO-US018590.
XX PR 08-JUN-2000; 2000US-0210759P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI White PS, Torney DC;
XX DR WPI; 2002-566450/60.
XX

```

PT Identifying sequences useful as address/capture tags for flow cytometry  
PT based minisequencing, by generating tag sequences and rejecting sequences  
PT based on certain parameters e.g. sequences which form stable hairpins.  
XX  
XX Disclosure; Page 9; 35pp; English.  
XX  
XX The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is a capture tag described in the  
CC exemplification of the invention  
XX  
XX Sequence 20 BP; 3 A; 4 C; 6 G; 7 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCTTTGGAGCGCTCTTAA 20  
DB 1 GGCTTTGGAGCGCTCTTAA 20  
RESULT 5  
AAS64515  
ID AAS64515 standard; cDNA; 545 BP.  
XX  
XX AC AAS64515;  
XX  
XX DT 13-FEB-2002 (first entry)  
XX  
XX DE DNA encoding novel human diagnostic protein #319.  
XX  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX PR 31-MAR-2000; 2000US-00540217.  
XX  
XX PR 23-AUG-2000; 2000US-00649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX DR WPI; 2001-639362/73.  
XX  
XX DR P-PSDB; ABG00328.  
XX  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX PS Claim 1; SEQ ID NO 319; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a

CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food imaging  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pat\_sequences  
XX  
XX Sequence 545 BP; 86 A; 164 C; 171 G; 124 T; 0 U; 0 Other;  
SQ  
Query Match 90.0%; Score 18; DB 5; Length 545;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCTTTGGAGCGCTCTTT 18  
DB 404 GGCTTTGGAGCGCTCTTT 421  
RESULT 6  
AAS64519  
ID AAS64519 standard; cDNA; 2076 BP.  
XX  
XX AC AAS64519;  
XX  
XX DT 13-FEB-2002 (first entry)  
XX  
XX DE DNA encoding novel human diagnostic protein #323.  
XX  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX PR 31-MAR-2000; 2000US-00540217.  
XX  
XX PR 23-AUG-2000; 2000US-00649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX DR WPI; 2001-639362/73.  
XX  
XX DR P-PSDB; ABG00332.  
XX  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX PS Claim 1; SEQ ID NO 323; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a

polypeptide in tissue, as molecular weight markers and as a food supplement. (ii) and its binding partners are useful in medical imaging of sites expressing (ii). (i) and (ii) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

Sequence 2076 BP; 465 A; 603 C; 596 G; 412 T; 0 U; 0 Other;  
 Query Match 90.0%; Score 18; DB 5; Length 2076;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTTGGAGCGCTCTTT 18  
 |||||  
 Db 538 GGCTTGGAGCGCTCTTT 555

RESULT 7  
 ID AAH41224/c  
 XX AAH41224 standard; DNA; 349980 BP.  
 AC AAH41224;  
 XX 29-OCT-2001 (first entry)  
 DT Pyrococcus abyssi genomic fragment #3.  
 DE Hyperthermophilic archaeon; hyperthermophilic protein; ds.  
 KW Pyrococcus abyssi.  
 OS Pyrococcus abyssi.  
 XX Location/Qualifiers  
 FH 1..49980  
 FT misc\_feature  
 FT /tag= a  
 FT /note= "This sequence overlaps with the 3' end of  
 AAH41223"  
 FT misc\_feature  
 FT 300001..349980  
 FT /tag= b  
 FT /note= "This sequence overlaps with the 5' end of  
 AAH41225"  
 XX  
 XX FR2792651-A1.  
 XX 27-OCT-2000.  
 XX 21-APR-1999; 99PR-00005034.  
 XX 21-APR-1999; 99PR-00005034.  
 XX (CNRS) CNRS CENT NAT RECH SCI.  
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.  
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 PI Querellou J, Weissenbach J, Saurin W, Heilig R;  
 XX WPI; 2001-126236/14.  
 XX New nucleotide sequences isolated from Pyrococcus abyssi encode proteins useful in industry.  
 XX Claim 1; Page 347-443; 1657pp; French.  
 XX The present invention relates to the genomic sequence of Pyrococcus abyssi and p. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal

vents. The present sequence is a fragment of the genomic sequence of P. abyssi. The 5' end of this sequence overlaps with the 3' end of AAH41223 and the 3' end of this sequence overlaps with the 5' end of AAH41225. The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to 110 degrees centigrade. Note: This patent is in the same patent family as WO200065062, which contains additional sequences as shown in AAB99132-AAB99143, AAH75903-AAH75920 and AAG66436

XX Sequence 349980 BP; 94090 A; 78692 C; 78319 G; 98879 T; 0 U; 0 Other;  
 Query Match 82.0%; Score 16.4; DB 5; Length 349980;  
 Best Local Similarity 94.4%; Pred. No. 2.6e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCTTGGAGCGCTCTTTA 19  
 |||||  
 Db 115453 GCTTGGAGCGCTCTTA 115436

RESULT 8  
 AAC18777  
 ID AAC18777 standard; cDNA; 258 BP.  
 XX AAC18777;  
 XX 06-OCT-2000 (first entry)  
 DT Human secreted protein 5' EST, SEQ ID NO: 22852.  
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN EP1033401-A2.  
 XX 06-SEP-2000.  
 XX 21-FEB-2000; 2000EP-00200510.  
 XX 26-FEB-1999; 99US-0122487P.  
 XX (GEST) GENSET.  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX WPI; 2000-500381/45.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX Claim 1; SEQ ID NO 22852; 71pp + Sequence Listing; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dr primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors  
 XX Sequence 258 BP; 68 A; 46 C; 65 G; 75 T; 0 U; 4 Other;  
 SQ Query Match 80.0%; Score 16; DB 3; Length 258;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTTAA 20  
||||| : : : : :  
Db 232 GGCTTTGARGGKCTCTTTAA 251

## RESULT 9

AAV27582/c  
ID AAV27582 standard; DNA; 1800 BP.

XX AAV27582;

XX 17-OCT-2003 (revised)

DT 12-OCT-1998 (first entry)

XX Helicobacter pylori flaB gene.

XX Flagellin; flaB gene; vaccine; immunogen; gastritis; peptic ulcer; ss.

XX Helicobacter pylori; strain CCUG 17874.

XX Key Location/Qualifiers

FT CDS 138..582

FT /\*tag= a

XX WO9823288-A1.

XX 04-JUN-1998.

XX 18-NOV-1997; 97WO-SE001928.

XX 25-NOV-1996; 96SE-00004322.

XX (ASTR ) ASTRA AB.

XX Boelin I, Berglindh T, Mellgard B, Svennerholm A;

XX WPI; 1998-322460/28.

XX P-PSDB; AAW61270.

XX Using Helicobacter pylori flagellin polypeptide as immunogen in vaccines  
PT - for treatment or prevention of Helicobacter pylori infection, provide  
PT strong and consistent immune response.

XX Disclosure; Page 26-29; 45pp; English.

XX This genomic DNA comprises the flaB gene of Helicobacter pylori CCUG

CC 17874 that codes for a flagellin polypeptide (see AAW61270). It was  
CC cloned from a H. pylori CCUG 17874 genomic library using probes obtained  
CC by PCR amplification of the 5' and 3' regions of the gene. The flagellin  
CC polypeptide FlaB or FlaA (see AAW61269), or its modified but  
CC antigenically equivalent forms, can be used to induce a protective immune  
CC response against H. pylori infection. Also claimed are vaccines

CC containing FlaA or FlaB and a carrier or diluent. The vaccines are used  
CC to treat or prevent H. pylori infection, particularly in humans.  
CC Flagellin is a strong and consistent antigen that stimulates a local  
CC immune response which decreases or eliminates colonisation of the gastric  
CC mucosa. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1800 BP; 545 A; 347 C; 448 G; 460 T; 0 U; 0 Other;

SQ Query Match 80.0%; Score 16; DB 2; Length 1800;

XX Best Local Similarity 100.0%; Pred. No. 2.1e+02;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCTTTGGAGCGCTCTT 17

||||| : : : : :  
Db 479 GGCTTTGGAGCGCTCTT 464

## RESULT 10

ADA49275/c  
ID ADA49275 standard; DNA; 582 BP.

XX ADA49275;

XX 20-NOV-2003 (first entry)

XX Maize gene conferring disease resistance in plants.

XX disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;  
XX maize.

XX Zea mays.

XX WO2003000906-A2.

XX 03-JAN-2003.

XX 21-JUN-2002; 2002WO-1B002453.

XX 22-JUN-2001; 2001US-0300112P.

PR 26-SEP-2001; 2001US-0352277P.

PR 22-MAR-2002; 2002US-0366535P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;

XX Katagiri F, Kreggs J, Provart N, Ricke D, Zhu T;

XX WPI; 2003-184052/18.

XX New polynucleotide comprising a plant nucleotide sequence having an open  
PT reading frame that encodes a polypeptide associated with disease  
PT resistance, useful for conferring resistance or tolerance to a plant  
PT pathogen.

XX Disclosure; SEQ ID NO 1345; 299pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a  
CC plant nucleotide sequence having an open reading frame that encodes a  
CC polypeptide associated with disease resistance or its fragment having  
CC substantially the same activity as the full-length polypeptide. The  
CC polynucleotide of the invention is useful for conferring resistance or  
CC tolerance to a plant pathogen. The present sequence represents a gene  
CC conferring disease resistance used in the invention.

XX Sequence 582 BP; 129 A; 141 C; 157 G; 155 T; 0 U; 0 Other;

XX Query Match 79.0%; Score 15.8; DB 8; Length 582;

XX Best Local Similarity 89.5%; Pred. No. 2.3e+02;

XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTTA 19

||||| : : : : :  
Db 530 GGCTTTGGATCGCTCTTCA 512

## RESULT 11

ABX66962

ID ABX66962 standard; DNA; 648 BP.

XX ABX66962;

XX 07-MAY-2003 (first entry)

XX Helicobacter pylori selected interacting domain (SID) DNA #1561.

XX Protein-protein interaction; ulcer; selected interacting domain; SID;

XX gene; ds.

XX Helicobacter pylori.

XX WO200266501-A2.





SQ Sequence 759 BP; 194 A; 183 C; 199 G; 183 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 7; Length 759;  
 Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCCTTGGAGCGCTCTTTAA 20  
 |||||  
 DB 696 GCTTTGGAGCGCACTTGAA 714

RESULT 14  
 ABX66344  
 ID ABX66344 standard; DNA; 1024 BP.  
 XX AC ABX66344;  
 XX DT 07-MAY-2003 (first entry)  
 DE Helicobacter pylori selected interacting domain (SID) DNA #943.  
 XX DE Protein-protein interaction; ulcer; selected interacting domain; SID;  
 XX KW gene; ds.  
 XX OS Helicobacter pylori.  
 XX OS WO200266501-A2.  
 XX EN 29-AUG-2002.  
 XX PD 28-DEC-2001; 2001WO-EP015428.  
 XX PF 02-JAN-2001; 2001US-0259302P.  
 XX PR (HYBR-) HYBRIGENICS.  
 XX PA (INSP) INST PASTEUR.  
 XX PI Legrain P, Rain J, Collard F, De Reuse H, Labigne A;  
 XX DR WPI; 2002-674910/72.  
 XX DR P-PSDB; ABU51600.  
 XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
 XX PT useful for identifying modulating compounds for treating or preventing  
 XX PT ulcers in mammals.  
 XX PS Claim 7; Page 314; 642pp; English.  
 XX CC The invention describes a complex of protein-protein interactions in  
 XX CC Helicobacter pylori selected from 421 complexes given in the  
 XX CC specification. The complex of protein-protein interactions are useful for  
 XX CC screening for agents which modulate the interaction of proteins.  
 XX CC Modulating compounds which binds to a targeted bacterial protein may be  
 XX CC used for treating or preventing ulcers in a human or animal. This  
 XX CC sequence encodes a selected interacting domain (SID), identified via  
 XX CC protein-protein interactions  
 XX SQ Sequence 1024 BP; 337 A; 179 C; 218 G; 290 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 6; Length 1024;  
 Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTTA 19  
 |||||  
 DB 59 GCCTTTGAAGCGCTCTATA 77

RESULT 15  
 ABX65416  
 ID ABX65416 standard; DNA; 1024 BP.  
 XX AC ABX65416;

XX 07-MAY-2003 (first entry)  
 DT Helicobacter pylori selected interacting domain (SID) DNA #15.  
 DE Protein-protein interaction; ulcer; selected interacting domain; SID;  
 XX KW gene; ds.  
 XX OS Helicobacter pylori.  
 XX OS WO200266501-A2.  
 XX EN 29-AUG-2002.  
 XX PD 28-DEC-2001; 2001WO-EP015428.  
 XX PF 02-JAN-2001; 2001US-0259302P.  
 XX PR (HYBR-) HYBRIGENICS.  
 XX PA (INSP) INST PASTEUR.  
 XX PI Legrain P, Rain J, Collard F, De Reuse H, Labigne A;  
 XX DR WPI; 2002-674910/72.  
 XX DR P-PSDB; ABU50672.  
 XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
 XX PT useful for identifying modulating compounds for treating or preventing  
 XX PT ulcers in mammals.  
 XX PS Claim 7; Page 87; 642pp; English.  
 XX CC The invention describes a complex of protein-protein interactions in  
 XX CC Helicobacter pylori selected from 421 complexes given in the  
 XX CC specification. The complex of protein-protein interactions are useful for  
 XX CC screening for agents which modulate the interaction of proteins.  
 XX CC Modulating compounds which binds to a targeted bacterial protein may be  
 XX CC used for treating or preventing ulcers in a human or animal. This  
 XX CC sequence encodes a selected interacting domain (SID), identified via  
 XX CC protein-protein interactions  
 XX SQ Sequence 1024 BP; 317 A; 206 C; 234 G; 267 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 6; Length 1024;  
 Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTTA 19  
 |||||  
 DB 749 GCCTTTGAAGCGCTCTATA 767

RESULT 16  
 ABL02097/c  
 ID ABL02097 standard; cDNA; 1061 BP.  
 XX AC ABL02097;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 773.  
 XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical; gene; ss.  
 XX OS Drosophila melanogaster.  
 XX OS WO200171042-A2.  
 XX PN 27-SEP-2001.  
 XX PD 23-MAR-2001; 2001WO-US009231.  
 XX PF  
 XX XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 ; Search time 27.1765 Seconds  
(without alignments)  
408.405 Million cell updates/sec

Title: US-09-877-819B-38

Perfect score: 20

Sequence: 1 ggccttgagcgctcttttaa 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	15.8	79.0	1851	4	US-09-134-000C-2975
2	15.8	79.0	4400	4	US-09-221-017B-995
3	15.2	76.0	1476	4	US-09-489-039A-792
C 4	15.2	76.0	3184	4	US-09-976-594-291
C 5	15.2	76.0	3411	2	US-08-849-480A-1
6	15.2	76.0	7152	3	US-09-167-681-29
7	15.2	76.0	8396	4	US-09-328-174A-1
8	15.2	76.0	8409	3	US-09-167-681-37
C 9	15.2	76.0	35100	2	US-08-770-379-18
C 10	15.2	76.0	35100	3	US-08-757-669A-18
C 11	15.2	76.0	35100	4	US-09-230-371A-18
C 12	15	75.0	66	4	US-09-523-656-36
C 13	15	75.0	105	4	US-09-523-656-33
C 14	15	75.0	4404	4	US-09-523-656-37
C 15	14.8	74.0	1299	4	US-09-489-039A-738
C 16	14.8	74.0	1982	4	US-09-221-017B-1068
C 17	14.8	74.0	2850	2	US-08-224-482-7
C 18	14.8	74.0	2922	4	US-09-489-039A-4404
C 19	14.8	74.0	3465	4	US-09-489-039A-553
C 20	14.8	74.0	3939	4	US-09-489-039A-4373
C 21	14.4	72.0	2907	4	US-09-620-312D-511
C 22	14.4	72.0	31208	4	US-09-852-067-3
C 23	14.2	71.0	58	3	US-09-140-466-10
C 24	14.2	71.0	463	4	US-09-280-116-73
C 25	14.2	71.0	499	4	US-09-889-914B-7
C 26	14.2	71.0	789	4	US-09-489-039A-4398
C 27	14.2	71.0	850	2	US-08-560-398-7

Sequence 4964, Ap  
Sequence 23, Appl  
Sequence 699, App  
Sequence 477, App  
Sequence 1296, Ap  
Sequence 1296, Ap  
Sequence 5098, Ap  
Sequence 8, Appli  
Sequence 1, Appli  
Sequence 4923, Ap  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 253, App  
Sequence 293, App  
Sequence 1, Appli  
Sequence 308, App  
Sequence 820, App  
Sequence 97, Appl  
Sequence 21, Appl  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 54, Appl  
Sequence 65, Appl  
Sequence 62, Appl  
Sequence 8, Appli  
Sequence 106, App  
Sequence 1603, Ap  
Sequence 21, Appl  
Sequence 137, App  
Sequence 73, Appl  
Sequence 1717, A  
Sequence 19139, A  
Sequence 19136, A  
Sequence 126, App  
Sequence 17716, A  
Sequence 1822, Ap  
Sequence 313, App  
Sequence 1016, App  
Sequence 192, App  
Sequence 192, App  
Sequence 192, App  
Sequence 192, App  
Sequence 12, Appl  
Sequence 284, App  
Sequence 2, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 1085, Ap  
Sequence 905, App  
Sequence 2558, Ap  
Sequence 253, App  
Sequence 2522, Ap  
Sequence 99, Appl  
Sequence 2627, Ap  
Sequence 9, Appli  
Sequence 1, Appli  
Sequence 959, App  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 516, App  
Sequence 3, Appli  
Sequence 1027, Ap

C 101 13.8 69.0 2856 1 US-09-041-075A-16  
 C 102 13.8 69.0 2856 3 US-08-716-873-3  
 C 103 13.8 69.0 2856 3 US-08-716-873-13  
 C 104 13.8 69.0 2856 3 US-08-716-873-14  
 C 105 13.8 69.0 2856 3 US-09-368-431-3  
 C 106 13.8 69.0 2856 3 US-09-368-431-13  
 C 107 13.8 69.0 2856 3 US-09-368-431-14  
 C 108 13.8 69.0 2877 1 US-09-041-075A-18  
 C 109 13.8 69.0 2885 4 US-09-016-434-1143  
 C 110 13.8 69.0 3078 4 US-09-711-164-272  
 C 111 13.8 69.0 3296 4 US-09-907-794A-310  
 C 112 13.8 69.0 3296 4 US-09-905-125A-310  
 C 113 13.8 69.0 3296 4 US-09-902-775A-310  
 C 114 13.8 69.0 3546 4 US-09-976-594-126  
 C 115 13.8 69.0 4027 2 US-08-551-356-5  
 C 116 13.8 69.0 4027 5 PCT-US93-12687-5  
 C 117 13.8 69.0 4140 3 US-08-716-873-1  
 C 118 13.8 69.0 4140 3 US-09-368-431-1  
 C 119 13.8 69.0 4978 1 US-08-220-603A-1  
 C 120 13.8 69.0 5852 4 US-08-833-768-10  
 C 121 13.8 69.0 6688 3 US-09-381-862-5  
 C 122 13.8 69.0 7037 4 US-09-853-768-3  
 C 123 13.8 69.0 7679 4 US-09-220-132-38  
 C 124 13.8 69.0 7680 4 US-09-023-655-1289  
 C 125 13.8 69.0 7680 5 PCT-US95-09819-6  
 C 126 13.8 69.0 7705 1 US-08-259-569-16  
 C 127 13.8 69.0 7705 2 US-08-826-885-16  
 C 128 13.8 69.0 7705 6 5455158-2  
 C 129 13.8 69.0 7803 2 US-08-551-356-1  
 C 130 13.8 69.0 7803 5 PCT-US93-12687-1  
 C 131 13.8 69.0 8044 4 US-09-566-921-135  
 C 132 13.8 69.0 8140 1 US-08-297-294A-1  
 C 133 13.8 69.0 11492 3 US-08-991-840A-1  
 C 134 13.8 69.0 14231 4 US-08-961-527-81  
 C 135 13.8 69.0 4403765 3 US-09-103-840A-2  
 C 136 13.8 69.0 4411529 3 US-09-103-840A-1  
 C 137 13.6 68.0 79 4 US-09-025-769B-113  
 C 138 13.6 68.0 79 4 US-08-025-769B-118  
 C 139 13.6 68.0 280 4 US-09-313-294A-2122  
 C 140 13.6 68.0 345 4 US-09-833-381-1863  
 C 141 13.6 68.0 351 4 US-09-220-132-81  
 C 142 13.6 68.0 361 4 US-09-025-769B-56  
 C 143 13.6 68.0 361 4 US-09-025-769B-58  
 C 144 13.6 68.0 390 4 US-09-621-976-10498  
 C 145 13.6 68.0 433 1 US-08-428-733A-1  
 C 146 13.6 68.0 444 4 US-09-489-039A-3299  
 C 147 13.6 68.0 465 4 US-09-328-352-3402  
 C 148 13.6 68.0 481 4 US-09-621-976-237  
 C 149 13.6 68.0 489 4 US-09-134-000C-461  
 C 150 13.6 68.0 539 6 5422263-3  
 C 151 13.6 68.0 594 4 US-09-134-000C-2075  
 C 152 13.6 68.0 610 3 US-09-385-982-303  
 C 153 13.6 68.0 716 3 US-09-328-111-701  
 C 154 13.6 68.0 774 4 US-09-489-039A-5576  
 C 155 13.6 68.0 905 4 US-09-976-594-788  
 C 156 13.6 68.0 930 4 US-09-540-236-107  
 C 157 13.6 68.0 1001 4 US-09-641-638-440  
 C 158 13.6 68.0 1001 4 US-09-641-638-453  
 C 159 13.6 68.0 1001 4 US-09-641-638-454  
 C 160 13.6 68.0 1260 4 US-09-166-350-22  
 C 161 13.6 68.0 1320 4 US-09-023-655-491  
 C 162 13.6 68.0 1404 4 US-09-173-151A-7  
 C 163 13.6 68.0 1495 4 US-09-016-434-1190  
 C 164 13.6 68.0 1495 4 US-09-023-655-1021  
 C 165 13.6 68.0 2091 4 US-09-328-352-3888  
 C 166 13.6 68.0 2156 1 US-08-012-988A-1  
 C 167 13.6 68.0 2156 4 US-09-023-655-1247  
 C 168 13.6 68.0 2314 4 US-09-173-151A-9  
 C 169 13.6 68.0 2585 3 US-09-008-697A-7  
 C 170 13.6 68.0 2681 4 US-09-621-502-1  
 C 171 13.6 68.0 2681 4 US-09-616-530A-6  
 C 172 13.6 68.0 2730 1 US-08-344-536-1  
 C 173 13.6 68.0 2730 3 US-08-920-562-1

Sequence 16, Appl  
 Sequence 3, Appl  
 Sequence 13, Appl  
 Sequence 14, Appl  
 Sequence 3, Appl  
 Sequence 13, Appl  
 Sequence 14, Appl  
 Sequence 18, Appl  
 Sequence 1143, Ap  
 Sequence 272, App  
 Sequence 310, App  
 Sequence 310, App  
 Sequence 310, App  
 Sequence 126, App  
 Sequence 5, Appl  
 Sequence 5, Appl  
 Sequence 1, Appl  
 Sequence 1, Appl  
 Sequence 1, Appl  
 Sequence 10, Appl  
 Sequence 5, Appl  
 Sequence 3, Appl  
 Sequence 38, Appl  
 Sequence 1289, Ap  
 Sequence 6, Appl  
 Sequence 16, Appl  
 Sequence 16, Appl  
 Patent No. 5455158  
 Sequence 1, Appl  
 Sequence 1, Appl  
 Sequence 135, App  
 Sequence 1, Appl  
 Sequence 1, Appl  
 Sequence 81, Appl  
 Sequence 113, App  
 Sequence 118, App  
 Sequence 2122, Ap  
 Sequence 1863, Ap  
 Sequence 81, Appl  
 Sequence 56, Appl  
 Sequence 58, Appl  
 Sequence 10498, A  
 Sequence 1, Appl  
 Sequence 3299, Ap  
 Sequence 3402, Ap  
 Sequence 237, App  
 Sequence 461, App  
 Patent No. 5422263  
 Sequence 2075, Ap  
 Sequence 303, App  
 Sequence 701, App  
 Sequence 5576, Ap  
 Sequence 788, App  
 Sequence 107, App  
 Sequence 440, App  
 Sequence 453, App  
 Sequence 454, App  
 Sequence 22, Appl  
 Sequence 491, App  
 Sequence 7, Appl  
 Sequence 1190, Ap  
 Sequence 1021, Ap  
 Sequence 3888, Ap  
 Sequence 1, Appl  
 Sequence 1247, Ap  
 Sequence 9, Appl  
 Sequence 7, Appl  
 Sequence 1, Appl  
 Sequence 6, Appl  
 Sequence 1, Appl  
 Sequence 1, Appl

Sequence 2, Appl  
 Sequence 461, App  
 Sequence 3, Appl  
 Sequence 7, Appl  
 Sequence 189, Appl  
 Sequence 1, Appl  
 Sequence 4, Appl

ALIGNMENTS  
 RESULT 1  
 US-09-134-000C-2975  
 ; Sequence 2975, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2975  
 ; LENGTH: 1851  
 ; TYPE: DNA  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-2975  
 Query Match 79.0%; Score 15.8; DB 4; Length 1851;  
 Best Local Similarity 89.5%; Pred. No. 30;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCCTTGAGCGCTCTTTAA 20  
 Db 280 GATTGGGCGCTTTTAA 298

RESULT 2  
 US-09-221-017B-995  
 ; Sequence 995, Application US/09221017B  
 ; Patent No. 6444799  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ross, Bruce C.  
 ; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 1120  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORRISON & FOERSTER  
 ; STREET: 755 PAGE MILL ROAD  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304-1018  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/221,017B  
 ; FILING DATE: 23-DEC-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PP1182  
 ; FILING DATE: 31-DEC-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PP1546  
 ; FILING DATE: 30-JAN-1998  
 ; PRIOR APPLICATION DATA:

RESULT 4  
US-09-976-594-291/c  
; Sequence 291, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael

ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,480A  
FILING DATE: 02-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/04711  
FILING DATE: 30-NOV-1995  
PRIOR APPLICATION DATA: DE P4442370.3  
FILING DATE: 02-DEC-1994  
APPLICATION NUMBER: DE 19505645.0  
FILING DATE: 18-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: AISENBERG, Irwin M.  
REGISTRATION NUMBER: 49,007  
REFERENCE/DOCKET NUMBER: 8125/P60984USO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202/638-6666  
TELEFAX: 202/393-5350  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3411 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Helicobacter pylori  
; STRAIN: Helicobacter pylori 69A  
; INDIVIDUAL ISOLATE: Clinical isolate 69A  
; IMMEDIATE SOURCE:  
; LIBRARY: Helicobacter pylori 69A - gene library in  
; LIBRARY: vector pRH160  
; CLONE: pRH439  
US-08-849-480A-1

Query Match 76.0%; Score 15.2; DB 2; Length 3411;  
Best Local Similarity 85.0%; Pred. No. 70;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
||||| ||||| |||||  
Db 2946 GGCTTTATAGCGTCTTTAA 2927

RESULT 6  
US-09-167-681-29  
; Sequence 29, Application US/09167681A  
; Patent No. 6265561  
; GENERAL INFORMATION:  
; APPLICANT: Weinshilboum, M.D., Richard M.  
; APPLICANT: Raftogianis, Rebecca B.  
; APPLICANT: Wood, Thomas C.  
; APPLICANT: Otterness, Diane M.  
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
; FILE REFERENCE: 07039/118001  
; CURRENT APPLICATION NUMBER: US/09167,681A  
; CURRENT FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 29  
; LENGTH: 7152  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3810)...(3956)  
; NAME/KEY: CDS  
; LOCATION: (4061)...(4186)  
; NAME/KEY: CDS  
; LOCATION: (4276)...(4374)  
; NAME/KEY: CDS  
; LOCATION: (5584)...(5709)  
; NAME/KEY: CDS  
; LOCATION: (5805)...(5900)  
; NAME/KEY: CDS  
; LOCATION: (6426)...(6605)  
; NAME/KEY: CDS  
; LOCATION: (6728)...(6837)  
US-09-167-681-29

Query Match 76.0%; Score 15.2; DB 3; Length 7152;  
Best Local Similarity 85.0%; Pred. No. 80;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
||||| ||||| |||||  
Db 2486 GGCTTTGAGATCACTTTAA 2505

RESULT 7  
US-09-328-174A-1  
; Sequence 1, Application US/09328174A  
; Patent No. 6448003  
; GENERAL INFORMATION:  
; APPLICANT: Guida, Marco

; APPLICANT: Kurth, Janice  
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase  
; TITLE OF INVENTION: (STP2)  
; FILE REFERENCE: 4389-6 (formerly SEQ-16P)  
; CURRENT APPLICATION NUMBER: US/09/328,174A  
; CURRENT FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: 09/328,174  
; PRIOR FILING DATE: 1999-06-08  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8396  
; TYPE: DNA  
; ORGANISM: H. sapiens  
US-09-328-174A-1

Query Match 76.0%; Score 15.2; DB 4; Length 8396;  
Best Local Similarity 85.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
||||| ||||| |||||  
Db 2417 GGCTTTGGAGATCACTTTAA 2436

RESULT 8  
US-09-167-681-37  
; Sequence 37, Application US/09167681A  
; Patent No. 6265561  
; GENERAL INFORMATION:  
; APPLICANT: Weinshilboum, M.D., Richard M.  
; APPLICANT: Raftogianis, Rebecca B.  
; APPLICANT: Wood, Thomas C.  
; APPLICANT: Otterness, Diane M.  
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
; FILE REFERENCE: 07039/118001  
; CURRENT APPLICATION NUMBER: US/09/167,681A  
; CURRENT FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 8397  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3730)...(3879)  
; NAME/KEY: CDS  
; LOCATION: (3987)...(4112)  
; NAME/KEY: CDS  
; LOCATION: (4198)...(4293)  
; NAME/KEY: CDS  
; LOCATION: (6288)...(6213)  
; NAME/KEY: CDS  
; LOCATION: (6309)...(6404)  
; NAME/KEY: CDS  
; LOCATION: (7214)...(7393)  
; NAME/KEY: CDS  
; LOCATION: (7316)...(7629)  
US-09-167-681-37

Query Match 76.0%; Score 15.2; DB 3; Length 8409;  
Best Local Similarity 85.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
||||| ||||| |||||  
Db 2418 GGCTTTGGAGATCACTTTAA 2437

RESULT 9  
US-08-770-379-18/c  
; Sequence 18, Application: US/08770379

```
/ Patent No. 5849564
/ GENERAL INFORMATION:
/ APPLICANT: Chang, Yuan
/ APPLICANT: Bohenzky, Roy A.
/ APPLICANT: Russo, James J.
/ APPLICANT: Edelman, Isidore S.
/ APPLICANT: Moore, Patrick S.
/ TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
/ TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ APPLICATION NUMBER: US/08/770,379
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 52342
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 18:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 35100 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-770-379-18

Query Match 76.0%; Score 15.2; DB 2; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTGAGCGCTCTTTAA 20
|||||
DB 23261 GGCCTTGAAGCGGCTTTAA 23242

RESULT 10
US-08-757-669A-18/c
/ Sequence 18, Application US/08/757669A
/ Patent No. 6193751
/ GENERAL INFORMATION:
/ APPLICANT: Chang, Yuan
/ APPLICANT: Bohenzky, Roy A.
/ APPLICANT: Russo, James J.
/ APPLICANT: Edelman, Isidore S.
/ APPLICANT: Moore, Patrick S.
/ TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
/ TITLE OF INVENTION: SEQUENCES AND USES THEREOF
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
```

```
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/757,669A
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 45185-F
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 18:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 35100 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-757-669A-18

Query Match 76.0%; Score 15.2; DB 3; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTGAGCGCTCTTTAA 20
|||||
DB 23261 GGCCTTGAAGCGGCTTTAA 23242

RESULT 11
US-09-230-371A-18/c
/ Sequence 18, Application US/09230371A
/ Patent No. 6348586
/ GENERAL INFORMATION:
/ APPLICANT: Chang, Yuan
/ APPLICANT: Bohenzky, Roy A.
/ APPLICANT: Russo, James J.
/ APPLICANT: Edelman, Isidore S.
/ APPLICANT: Moore, Patrick S.
/ TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: 45185-G-PCT-US
/ CURRENT APPLICATION NUMBER: US/09/230,371A
/ CURRENT FILING DATE: 1999-11-17
/ PRIOR APPLICATION NUMBER: PCT/US97/13346
/ PRIOR FILING DATE: 1997-07-22
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 18
/ LENGTH: 35100
/ TYPE: DNA
/ ORGANISM: Kaposi's sarcoma-associated herpesvirus
/ US-09-230-371A-18

Query Match 76.0%; Score 15.2; DB 4; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTGAGCGCTCTTTAA 20
|||||
DB 23261 GGCCTTGAAGCGGCTTTAA 23242

RESULT 12
US-09-523-656-36/c
/ Sequence 36, Application US/09523656
/ Patent No. 6458563
/ GENERAL INFORMATION:
/ APPLICANT: Lollar S., John
/ TITLE OF INVENTION: MODIFIED FACTOR VIII
```

; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; CURRENT FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707  
; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide  
; OTHER INFORMATION: primer  
US-09-523-656-36

Query Match 75.0%; Score 15; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTC 15  
| | | | | | | | | | | | | | | | | |  
Db 23 GGCTTTGGAGCGCTC 9

RESULT 13  
US-09-523-656-33/c  
; Sequence 33, Application US/09523656  
; Patent No. 6458563  
; GENERAL INFORMATION:  
; APPLICANT: Lollar S., John  
; TITLE OF INVENTION: MODIFIED FACTOR VIII  
; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; CURRENT FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707  
; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 33  
; LENGTH: 105  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: linker  
US-09-523-656-33

Query Match 75.0%; Score 15; DB 4; Length 105;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTC 15  
| | | | | | | | | | | | | | | | | |  
Db 62 GGCTTTGGAGCGCTC 48

RESULT 14  
US-09-523-656-37/c  
; Sequence 37, Application US/09523656  
; Patent No. 6458563  
; GENERAL INFORMATION:  
; APPLICANT: Lollar S., John  
; TITLE OF INVENTION: MODIFIED FACTOR VIII  
; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; CURRENT FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707

; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 37  
; LENGTH: 4404  
; TYPE: DNA  
; ORGANISM: Porcine  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: (1)..(4401)  
US-09-523-656-37

Query Match 75.0%; Score 15; DB 4; Length 4404;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTC 15  
| | | | | | | | | | | | | | | | | |  
Db 2324 GGCTTTGGAGCGCTC 2310

RESULT 15  
US-09-489-039A-738  
; Sequence 738, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 738  
; LENGTH: 1299  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-738

Query Match 74.0%; Score 14.8; DB 4; Length 1299;  
Best Local Similarity 88.9%; Pred. No. 97;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTTTGAGCGCTCTTTA 19  
| | | | | | | | | | | | | | | | | |  
Db 842 GCTTTGAGCGCTCTTTA 859

RESULT 16  
US-09-221-017B-1068/c  
; Sequence 1068, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998



```
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: P21182
/ FILING DATE: 31-DEC-1997
/ PRIOR APPLICATION DATA: P1546
/ APPLICATION NUMBER: P1546
/ FILING DATE: 30-JAN-1998
/ PRIOR APPLICATION DATA: P22911
/ APPLICATION NUMBER: P22911
/ FILING DATE: 09-APR-1998
/ PRIOR APPLICATION DATA: PCT/US98/01023
/ APPLICATION NUMBER: PCT/US98/01023
/ FILING DATE: 10-DEC-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Monroy, Gladys H
/ REGISTRATION NUMBER: 32,430
/ REFERENCE/DOCKET NUMBER: 27340-20021.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-813-5600
/ TELEFAX: 650-494-0792
/ TELEX: 706141
/ INFORMATION FOR SEQ ID NO: 1068:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1982 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: UNKNOWN
/ ORIGINAL SOURCE:
/ ORGANISM: PORPHYROMONAS GINGIVALIS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1...1982
/
/ US-09-221-017B-1068
/
/ Query Match 74.0%; Score 14.8; DB 4; Length 1982;
/ Best Local Similarity 88.9%; Pred. No. 1.e+02;
/ Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 1 GGCTTTGGAGCGCTCTTT 18
/ | | | | | | | | | | | | | | | |
/ Db 1666 GGCTTTGGAGCGCTCTTT 1649
/
/ RESULT 17
/ US-08-224-482-7/c
/ Sequence 7, Application US/08224482
/ Patent No. 5837592
/ GENERAL INFORMATION:
/ APPLICANT: Mercola, Dan
/ APPLICANT: Adamson, Eileen D.
/ TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
/ TITLE OF INVENTION: PDGF by Mammalian BGR
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Campbell and Flores
/ STREET: 4370 La Jolla Village Drive, Suite 700
/ CITY: San Diego
/ STATE: California
/ COUNTRY: USA
/ ZIP: 92122
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/224,482
/ FILING DATE: 07-APR-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/DOCKET NUMBER: P-ME 9913
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 535-9001
/ TELEFAX: (619) 535-8949
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2850 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 358..1519
/ US-08-224-482-7
/
/ Query Match 74.0%; Score 14.8; DB 2; Length 2850;
/ Best Local Similarity 88.9%; Pred. No. 1.1e+02;
/ Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 2 GCTTTGGAGCGCTCTTTA 19
/ | | | | | | | | | | | | | | |
/ Db 1966 GCTTTGGAGCGCTCTTTA 1949
/
/ RESULT 18
/ US-09-489-039A-4404/c
/ Sequence 4404, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 4404
/ LENGTH: 2922
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
/ US-09-489-039A-4404
/
/ Query Match 74.0%; Score 14.8; DB 4; Length 2922;
/ Best Local Similarity 88.9%; Pred. No. 1.1e+02;
/ Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 2 GCTTTGGAGCGCTCTTTA 19
/ | | | | | | | | | | | | | | |
/ Db 71 GCTTTGAAGCGCGCTTTA 54
/
/ RESULT 19
/ US-09-489-039A-553/c
/ Sequence 553, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 553
/ LENGTH: 3465
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
```

US-09-489-039A-553

Query Match 74.0%; Score 14.8; DB 4; Length 3465;  
Best Local Similarity 88.9%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTTTGGAGCGCTCTTTA 19  
|||||  
DB 2249 GCTTTGATGCGCTCTTTA 2232

RESULT 20

US-09-489-039A-4373  
Sequence 4373, Application US/09489039A

Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709-2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 4373

LENGTH: 3939

TYPE: DNA

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-4373

Query Match 74.0%; Score 14.8; DB 4; Length 3939;

Best Local Similarity 88.9%; Pred. No. 1.2e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTTTGGAGCGCTCTTTA 19

|||||

DB 3542 GCTTTGAAGCGCGCTTTA 3559

RESULT 21

US-09-620-312D-511

Sequence 511, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Tonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillingshast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: Pt\_Fl\_genes Version 1.0

SEQ ID NO 511  
LENGTH: 2907  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (359)..(2530)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2907)  
OTHER INFORMATION: n = a,t,c or g  
US-09-620-312D-511

Query Match 72.0%; Score 14.4; DB 4; Length 2907;  
Best Local Similarity 93.8%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCTTGGAGCGCTCT 16  
|||||  
DB 781 GCCTTGGTGGCTCT 796

RESULT 22

US-09-852-067-3/c

Sequence 3, Application US/09852067

Patent No. 6531297

GENERAL INFORMATION:

APPLICANT: MERKULOV, Gennady et al

TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,

TITLE OF INVENTION: AND USES THEREOF

FILE REFERENCE: C1000897-CIP

CURRENT APPLICATION NUMBER: US/09/852,067

CURRENT FILING DATE: 2001-05-10

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 31208

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(31208)

OTHER INFORMATION: n = A,T,C or G

US-09-852-067-3

Query Match 72.0%; Score 14.4; DB 4; Length 31208;

Best Local Similarity 93.8%; Pred. No. 2.8e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TTGGAGCGCTCTTTA 19

|||||

DB 18183 TTGGAGCTCTCTTTA 18168

RESULT 23

US-09-140-466-10/c

Sequence 10, Application US/09140466

Patent No. 6268160

GENERAL INFORMATION:

APPLICANT: CLOUGH, BARBARA

APPLICANT: PREISER, PETER

APPLICANT: WILSON, ROBERT

TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE

TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS

TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS

FILE REFERENCE: N68837B GCW PJC DP

CURRENT APPLICATION NUMBER: US/09/140,466

CURRENT FILING DATE: 1998-08-26

EARLIER APPLICATION NUMBER: US 60/056,246

EARLIER FILING DATE: 1997-08-28

NUMBER OF SEQ ID NOS: 14

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:48:43 ; Search time 122.706 Seconds  
(without alignments)  
729.318 Million cell updates/sec

Title: US-09-877-819B-38

Perfect score: 20

Sequence: 1 ggcttggagcgtctttaa 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 80 summaries

#### Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	10	US-09-877-819B-3
2	20	100.0	20	10	US-09-877-819B-4
3	20	100.0	20	10	US-09-877-819B-36
4	20	100.0	20	10	US-09-877-819B-38
5	16.8	84.0	494	13	US-10-027-632-3532
6	16.8	84.0	494	16	US-10-027-632-3532
7	16.8	84.0	1018	13	US-10-425-114-10111
8	16.8	84.0	1299	13	US-10-425-114-20063
9	16.8	84.0	2382	13	US-10-424-599-132396
10	16.4	82.0	434	13	US-10-424-599-79300
11	16.4	80.0	1800	8	US-08-973-028-3
12	15.8	79.0	790	13	US-10-424-599-6560
13	15.8	79.0	820	13	US-10-027-632-173218
14	15.8	79.0	820	16	US-10-027-632-173218

9	US-09-815-242-7264	79.0	1110	1110	Sequence 7264, Ap
13	US-10-282-122A-22599	79.0	1110	1110	Sequence 22599, A
13	US-09-882-227-89	79.0	1170	1170	Sequence 89, Appl
13	US-10-282-122A-39670	79.0	1581	1581	Sequence 39670, A
13	US-10-194-163-995	79.0	4400	4400	Sequence 995, App
9	US-09-070-927A-146	79.0	8546	8546	Sequence 146, App
16	US-10-401-194-1	79.0	68571	68571	Sequence 1, Appli
9	US-09-738-626-1	79.0	3309400	3309400	Sequence 1, Appli
13	US-10-335-977-3971	77.0	732	732	Sequence 3971, Ap
13	US-10-335-977-3970	77.0	735	735	Sequence 3970, Ap
14	US-10-001-873-17	77.0	1176	1176	Sequence 17, Appl
9	US-09-895-913A-149	77.0	1440	1440	Sequence 149, App
10	US-09-764-872-732	77.0	10766	10766	Sequence 792, App
9	US-09-867-701-2971	76.0	187	187	Sequence 2971, Ap
13	US-10-424-599-41104	76.0	236	236	Sequence 41104, A
13	US-10-085-783A-1558	76.0	236	236	Sequence 1558, Ap
13	US-10-242-535A-1558	76.0	253	253	Sequence 1558, Ap
9	US-09-923-876-327	76.0	265	265	Sequence 327, App
11	US-09-923-876-327	76.0	265	265	Sequence 327, App
13	US-10-335-977-1774	76.0	270	270	Sequence 1774, Ap
13	US-10-335-977-1773	76.0	273	273	Sequence 1773, Ap
13	US-10-424-599-513	76.0	366	366	Sequence 513, App
10	US-09-918-993-5244	76.0	444	444	Sequence 5244, Ap
13	US-10-424-599-72018	76.0	456	456	Sequence 72018, A
13	US-10-424-599-72018	76.0	526	526	Sequence 141450, A
13	US-10-027-632-259306	76.0	555	555	Sequence 259306, A
13	US-10-027-632-259307	76.0	555	555	Sequence 259307, A
16	US-10-027-632-259306	76.0	555	555	Sequence 259306, A
16	US-10-027-632-259307	76.0	555	555	Sequence 259307, A
15	US-10-029-386-9951	76.0	570	570	Sequence 9951, Ap
15	US-10-335-977-2171	76.0	615	615	Sequence 2171, Ap
9	US-09-738-626-3026	76.0	747	747	Sequence 3026, Ap
9	US-09-974-300-1956	76.0	762	762	Sequence 1956, Ap
9	US-10-335-977-416	76.0	765	765	Sequence 416, App
13	US-10-335-977-415	76.0	768	768	Sequence 415, App
13	US-10-424-599-105581	76.0	787	787	Sequence 105581, A
13	US-10-335-977-1776	76.0	807	807	Sequence 1776, Ap
16	US-10-369-493-46752	76.0	948	948	Sequence 46752, A
9	US-09-815-242-7432	76.0	1110	1110	Sequence 7432, Ap
13	US-10-335-977-782	76.0	1287	1287	Sequence 782, App
13	US-10-424-599-105582	76.0	1576	1576	Sequence 105582, A
13	US-10-424-599-69421	76.0	1640	1640	Sequence 69421, A
16	US-10-369-493-33650	76.0	1876	1876	Sequence 33650, A
16	US-10-369-493-35696	76.0	2052	2052	Sequence 35696, A
9	US-09-815-242-7440	76.0	2061	2061	Sequence 7440, Ap
9	US-09-738-626-3025	76.0	2253	2253	Sequence 3025, Ap
13	US-10-027-632-112010	76.0	2385	2385	Sequence 112010, A
16	US-10-027-632-112010	76.0	2385	2385	Sequence 112010, A
16	US-10-108-260A-2360	76.0	2550	2550	Sequence 2360, Ap
13	US-09-795-651-49	76.0	3420	3420	Sequence 49, Appl
13	US-10-424-599-105261	76.0	5945	5945	Sequence 105261, A
13	US-09-829-545-29	76.0	7152	7152	Sequence 29, Appl
15	US-10-206-839-1	76.0	8396	8396	Sequence 1, Appli
13	US-09-829-545-37	76.0	8397	8397	Sequence 37, Appl
13	US-10-087-192-1663	76.0	193853	193853	Sequence 1663, Ap
15	US-10-175-523-58	76.0	251364	251364	Sequence 58, Appl
15	US-10-175-523-61	76.0	251364	251364	Sequence 61, Appl
15	US-10-175-523-79	76.0	251364	251364	Sequence 79, Appl
13	US-10-027-632-123862	75.0	1238	1238	Sequence 123862, A
16	US-10-027-632-123862	75.0	1238	1238	Sequence 123862, A
13	US-10-424-599-98627	74.0	267	267	Sequence 98627, A
9	US-09-974-300-8240	74.0	301	301	Sequence 8240, Ap
9	US-09-864-761-10521	74.0	352	352	Sequence 10521, A
16	US-10-259-194A-153	74.0	360	360	Sequence 153, App
9	US-09-864-761-24452	74.0	386	386	Sequence 24452, A
9	US-09-864-761-27143	74.0	408	408	Sequence 27143, A
9	US-09-772-134B-51	74.0	480	480	Sequence 51, Appl
9	US-09-772-134B-52	74.0	480	480	Sequence 52, Appl
9	US-09-864-761-7758	74.0	534	534	Sequence 7758, Ap
13	US-10-027-632-123839	74.0	583	583	Sequence 123839, A
13	US-10-027-632-125840	74.0	583	583	Sequence 125840, A
13	US-10-027-632-125841	74.0	583	583	Sequence 125841, A
16	US-10-027-632-125839	74.0	583	583	Sequence 125839, A

```

86 14.8 74.0 583 16 US-10-027-632-125840 Sequence 125840, 13 US-10-335-977-2504 Sequence 2504, Ap
87 14.8 74.0 583 16 US-10-027-632-125841 Sequence 125841, 16 US-10-104-047-635 Sequence 635, App
88 14.8 74.0 583 16 US-10-027-632-89299 Sequence 89299, A c 163 16 US-10-282-122A-24913 Sequence 24913, A
89 14.8 74.0 605 13 US-10-027-632-89299 Sequence 89299, A c 163 16 US-10-282-122A-24913 Sequence 24913, A
90 14.8 74.0 605 13 US-10-027-632-89300 Sequence 89300, A c 164 16 US-09-764-891-9475 Sequence 9475, Ap
91 14.8 74.0 605 13 US-10-027-632-89300 Sequence 89300, A c 164 16 US-09-764-891-9475 Sequence 9475, Ap
92 14.8 74.0 605 13 US-10-027-632-317126 Sequence 317126, c 165 16 US-10-205-428-856 Sequence 856, App
93 14.8 74.0 605 13 US-10-027-632-317127 Sequence 317127, c 166 16 US-10-205-428-856 Sequence 856, App
94 14.8 74.0 605 13 US-10-027-632-317128 Sequence 317128, c 167 16 US-10-205-428-857 Sequence 857, App
95 14.8 74.0 605 13 US-10-027-632-89299 Sequence 89299, A c 168 16 US-10-037-270-511 Sequence 511, App
96 14.8 74.0 605 16 US-10-027-632-89300 Sequence 89300, A c 169 16 US-10-117-722-511 Sequence 511, App
97 14.8 74.0 605 16 US-10-027-632-317126 Sequence 317126, c 170 16 US-10-017-161-709 Sequence 709, App
98 14.8 74.0 605 16 US-10-027-632-317127 Sequence 317127, c 171 16 US-10-232-798-621 Sequence 621, App
99 14.8 74.0 605 16 US-10-027-632-317128 Sequence 317128, c 172 16 US-10-232-798-621 Sequence 621, App
100 14.8 74.0 746 13 US-10-424-599-70431 Sequence 70431, A c 173 16 US-09-852-067-3 Sequence 3, Appli
101 14.8 74.0 765 9 US-09-738-626-75 Sequence 75, Appli c 174 16 US-10-338-691-3 Sequence 1147, Ap
102 14.8 74.0 777 16 US-10-398-221-741 Sequence 741, App c 175 16 US-10-087-192-1147 Sequence 107, Appli
103 14.8 74.0 818 13 US-10-424-599-63507 Sequence 63507, A c 176 16 US-10-002-491-10 Sequence 1, Appli
104 14.8 74.0 972 13 US-10-027-632-9130 Sequence 9130, Ap c 177 16 US-10-067-514-1 Sequence 1, Appli
105 14.8 74.0 972 16 US-10-027-632-9130 Sequence 9130, Ap c 177 16 US-10-067-514-1 Sequence 1, Appli
106 14.8 74.0 1050 15 US-10-239-431A-3 Sequence 5, Appli c 178 16 US-09-845-335-10 Sequence 10, Appli
107 14.8 74.0 1128 15 US-10-239-431A-3 Sequence 3, Appli c 179 16 US-10-029-386-22085 Sequence 22085, A
108 14.8 74.0 1156 9 US-09-917-800A-1347 Sequence 1347, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
109 14.8 74.0 1156 15 US-10-205-134-32 Sequence 32, Appl c 180 16 US-09-864-761-21439 Sequence 21439, A
110 14.8 74.0 1156 16 US-10-388-934-159 Sequence 159, App c 180 16 US-09-864-761-21439 Sequence 21439, A
111 14.8 74.0 1467 9 US-09-738-626-76 Sequence 76, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
112 14.8 74.0 1830 15 US-10-239-431A-1 Sequence 1, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
113 14.8 74.0 1920 16 US-10-398-221-743 Sequence 743, App c 180 16 US-09-864-761-21439 Sequence 21439, A
114 14.8 74.0 1920 16 US-10-398-221-743 Sequence 743, App c 180 16 US-09-864-761-21439 Sequence 21439, A
115 14.8 74.0 1920 16 US-10-398-221-743 Sequence 743, App c 180 16 US-09-864-761-21439 Sequence 21439, A
116 14.8 74.0 1982 13 US-10-194-163-1068 Sequence 1068, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
117 14.8 74.0 2610 16 US-10-369-493-37578 Sequence 37578, A c 180 16 US-09-864-761-21439 Sequence 21439, A
118 14.8 74.0 2631 13 US-10-282-122A-41142 Sequence 41142, A c 180 16 US-09-864-761-21439 Sequence 21439, A
119 14.8 74.0 2726 9 US-09-764-864-772 Sequence 772, App c 180 16 US-09-864-761-21439 Sequence 21439, A
120 14.8 74.0 2744 9 US-09-764-864-357 Sequence 357, App c 180 16 US-09-864-761-21439 Sequence 21439, A
121 14.8 74.0 3074 13 US-10-282-122A-19458 Sequence 19458, A c 180 16 US-09-864-761-21439 Sequence 21439, A
122 14.8 74.0 3444 13 US-10-282-122A-23365 Sequence 23365, A c 180 16 US-09-864-761-21439 Sequence 21439, A
123 14.8 74.0 3872 13 US-10-282-122A-23663 Sequence 23663, A c 180 16 US-09-864-761-21439 Sequence 21439, A
124 14.8 74.0 4285 10 US-09-960-706-1074 Sequence 1074, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
125 14.8 74.0 4338 15 US-10-247-671-104 Sequence 104, App c 180 16 US-09-864-761-21439 Sequence 21439, A
126 14.8 74.0 4339 14 US-10-002-600-50 Sequence 50, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
127 14.8 74.0 148831 9 US-09-969-708-79 Sequence 79, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
128 14.8 74.0 148831 9 US-09-954-456-2116 Sequence 2116, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
129 14.8 74.0 148831 9 US-09-873-367C-646 Sequence 646, App c 180 16 US-09-864-761-21439 Sequence 21439, A
130 14.8 74.0 148831 13 US-10-240-425-363 Sequence 363, App c 180 16 US-09-864-761-21439 Sequence 21439, A
131 14.8 74.0 653122 13 US-10-087-192-226 Sequence 226, App c 180 16 US-09-864-761-21439 Sequence 21439, A
132 14.8 74.0 3309400 9 US-09-738-626-1 Sequence 1, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
133 14.8 74.0 426 9 US-09-864-761-21627 Sequence 21627, A c 180 16 US-09-864-761-21439 Sequence 21439, A
134 14.4 72.0 186 13 US-10-282-122A-11502 Sequence 11502, A c 180 16 US-09-864-761-21439 Sequence 21439, A
135 14.4 72.0 195 13 US-10-282-122A-14244 Sequence 14244, A c 180 16 US-09-864-761-21439 Sequence 21439, A
136 14.4 72.0 213 9 US-09-864-761-21445 Sequence 21445, A c 180 16 US-09-864-761-21439 Sequence 21439, A
137 14.4 72.0 268 13 US-10-424-599-72721 Sequence 72721, A c 180 16 US-09-864-761-21439 Sequence 21439, A
138 14.4 72.0 309 13 US-10-085-783A-56225 Sequence 56225, A c 180 16 US-09-864-761-21439 Sequence 21439, A
139 14.4 72.0 309 15 US-10-242-635A-56225 Sequence 56225, A c 180 16 US-09-864-761-21439 Sequence 21439, A
140 14.4 72.0 426 9 US-09-864-761-4705 Sequence 4705, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
141 14.4 72.0 435 9 US-09-864-761-4893 Sequence 4893, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
142 14.4 72.0 474 10 US-09-918-895-29217 Sequence 29217, A c 180 16 US-09-864-761-21439 Sequence 21439, A
143 14.4 72.0 504 13 US-10-243-552-224 Sequence 224, App c 180 16 US-09-864-761-21439 Sequence 21439, A
144 14.4 72.0 511 13 US-10-424-599-93201 Sequence 93201, A c 180 16 US-09-864-761-21439 Sequence 21439, A
145 14.4 72.0 536 16 US-10-341-961A-253 Sequence 253, App c 180 16 US-09-864-761-21439 Sequence 21439, A
146 14.4 72.0 581 13 US-10-027-632-15230 Sequence 15230, A c 180 16 US-09-864-761-21439 Sequence 21439, A
147 14.4 72.0 581 16 US-10-027-632-15230 Sequence 15230, A c 180 16 US-09-864-761-21439 Sequence 21439, A
148 14.4 72.0 609 13 US-10-335-977-3577 Sequence 3577, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
149 14.4 72.0 611 13 US-10-027-632-229898 Sequence 229898, c 180 16 US-09-864-761-21439 Sequence 21439, A
150 14.4 72.0 611 16 US-10-027-632-229898 Sequence 229898, c 180 16 US-09-864-761-21439 Sequence 21439, A
151 14.4 72.0 663 9 US-09-841-132-409 Sequence 409, App c 180 16 US-09-864-761-21439 Sequence 21439, A
152 14.4 72.0 1251 13 US-10-335-977-1149 Sequence 1149, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
153 14.4 72.0 1266 13 US-10-335-977-1150 Sequence 1150, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
154 14.4 72.0 1278 13 US-10-335-977-1151 Sequence 1151, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
155 14.4 72.0 1508 13 US-10-302-172-109 Sequence 109, App c 180 16 US-09-864-761-21439 Sequence 21439, A
156 14.4 72.0 2118 13 US-10-282-122A-41372 Sequence 41372, A c 180 16 US-09-864-761-21439 Sequence 21439, A
157 14.4 72.0 2259 13 US-10-335-977-2502 Sequence 2502, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A
158 14.4 72.0 2456 13 US-09-939-853A-34 Sequence 34, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
159 14.4 72.0 2456 13 US-09-939-853A-36 Sequence 36, Appli c 180 16 US-09-864-761-21439 Sequence 21439, A
160 14.4 72.0 2547 13 US-10-335-977-2503 Sequence 2503, Ap c 180 16 US-09-864-761-21439 Sequence 21439, A

```

## ALIGNMENTS

## RESULT 1

```

; Sequence 3, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address tag
US-09-877-819B-3

```

```

Query Match 100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGCTTTGGAGCGCTCTTAA 20
|||||
DB 1 GGCTTTGGAGCGCTCTTAA 20
|||||

```

## RESULT 2

```

; Sequence 4, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

```

; OTHER INFORMATION: Capture tag  
US-09-877-819B-4

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 20 GGCTTTGGAGCGCTCTTTAA 1

## RESULT 3

US-09-877-819B-36/c  
; Sequence 36, Application US/09877819B  
; Publication No. US20030190609A1  
; GENERAL INFORMATION:  
; APPLICANT: Torney, David  
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing  
; FILE REFERENCE: S-94,664  
; CURRENT APPLICATION NUMBER: US/09/877,819B  
; CURRENT FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 36  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Capture sequence  
US-09-877-819B-36

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 20 GGCTTTGGAGCGCTCTTTAA 1

## RESULT 4

US-09-877-819B-38  
; Sequence 38, Application US/09877819B  
; Publication No. US20030190609A1  
; GENERAL INFORMATION:  
; APPLICANT: Torney, David  
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing  
; FILE REFERENCE: S-94,664  
; CURRENT APPLICATION NUMBER: US/09/877,819B  
; CURRENT FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 38  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Address sequence  
US-09-877-819B-38

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 1 GGCTTTGGAGCGCTCTTTAA 20

## RESULT 5

US-10-027-632-3532  
; Sequence 3532, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3532  
; LENGTH: 494  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-3532

Query Match 84.0%; Score 16.8; DB 13; Length 494;  
Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 315 GGCTTTGGAGCACTCTTTGA 334

## RESULT 6

US-10-027-632-3532  
; Sequence 3532, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3532  
; LENGTH: 494  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-3532

Query Match 84.0%; Score 16.8; DB 16; Length 494;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
 ||||| ||||| ||||| |||||  
 Db 315 GGCTTTGGAGCACTCTTTGA 334

RESULT 7  
 US-10-425-114-10111  
 ; Sequence 10111, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 10111  
 ; LENGTH: 1018  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURES:  
 ; OTHER INFORMATION: Clone ID: 700894185\_FLI  
 US-10-425-114-10111

Query Match 84.0%; Score 16.8; DB 13; Length 1018;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
 ||||| ||||| ||||| |||||  
 Db 835 GGCTTTGAAAGCTCTTTAA 854

RESULT 8  
 US-10-425-114-20063  
 ; Sequence 20063, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 20063  
 ; LENGTH: 1299  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURES:  
 ; OTHER INFORMATION: Clone ID: LIB3170-053-F12\_FLI  
 US-10-425-114-20063

Query Match 84.0%; Score 16.8; DB 13; Length 1299;  
 Best Local Similarity 90.0%; Pred. No. 1.3e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
 ||||| ||||| ||||| |||||  
 Db 498 GGCTTTGAAGCTCTCTTTAA 517

RESULT 9  
 US-10-424-599-132396  
 ; Sequence 132396, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 132396  
 ; LENGTH: 2382  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURES:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90560C.1  
 US-10-424-599-132396

Query Match 84.0%; Score 16.8; DB 13; Length 2382;  
 Best Local Similarity 90.0%; Pred. No. 1.3e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCTTTGGAGCGCTCTTTAA 20  
 ||||| ||||| ||||| |||||  
 Db 2131 GGCTTTGAAAGCTCTTTAA 2150

RESULT 10  
 US-10-424-599-79300/c  
 ; Sequence 79300, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 79300  
 ; LENGTH: 434  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURES:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_42623C.1  
 US-10-424-599-79300

Query Match 82.0%; Score 16.4; DB 13; Length 434;  
 Best Local Similarity 94.4%; Pred. No. 1.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCCTTTGGAGCGCTCTTTA 19  
 ||||| ||||| ||||| |||||  
 Db 357 GCCTTTGGAGCGCTCTTTA 340

RESULT 11  
 US-08-973-028-3/c  
 ; Sequence 3, Application US/08973028  
 ; Publication No. US2002028210A1

GENERAL INFORMATION:  
APPLICANT: Berglindh, Thomas  
APPLICANT: Bolin, Ingrid  
APPLICANT: Mellgard, Bjorn  
APPLICANT: Svennerholm, Ann-Mari  
TITLE OF INVENTION: A Vaccine Composition Comprising Helicobacter  
TITLE OF INVENTION: Pylori Flagellin Polypeptide  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,028  
FILING DATE: 03-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE97/01928  
FILING DATE: 18-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9604322-9  
FILING DATE: 25-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cleland, Thelma A. Chen  
REGISTRATION NUMBER: 40,948  
REFERENCE/DOCKET NUMBER: 1103326-0289  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8515  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1800 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 138..1682  
OTHER INFORMATION: /product= "FlaB protein"

Query Match 80.0%; Score 16; DB 8; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Query 2 GCCTTTGGAGCGCTCTTT 17  
Db 479 GCCTTTGGAGCGCTCTTT 464  
RESULT 12  
US-10-424-599-6560  
Sequence 6560, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 6560  
LENGTH: 790  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_105933C.1  
US-10-424-599-6560  
Query Match 79.0%; Score 15.8; DB 13; Length 790;  
Best Local Similarity 89.5%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Query 1 GGCTTTGGAGCGCTCTTTTA 19  
Db 603 GGCTTTGGAGCGCTCTTTTA 621  
RESULT 13  
US-10-027-632-173218  
Sequence 173218, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,219  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 173218  
LENGTH: 820  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-173218  
Query Match 79.0%; Score 15.8; DB 13; Length 820;  
Best Local Similarity 89.5%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Query 1 GGCTTTGGAGCGCTCTTTTA 19  
Db 492 GGCTTTGGAGCGCTCTTTTA 510  
RESULT 14  
US-10-027-632-173218  
Sequence 173218, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 173218  
; LENGTH: 820  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-173218

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 16; Length 820;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCATTGAGCGCTCTTTA 19  
||||| ||||| ||||| |||||  
Db 492 GGCATTGAGCGCTCTTTA 510

RESULT 15  
US-09-815-242-7264  
; Sequence 7264, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7264  
; LENGTH: 1110  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1110)  
US-09-815-242-7264

Query Match  
79.0%; Score 15.8; DB 9; Length 1110;

Best Local Similarity 89.5%; Pred. No. 4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCATTGAGCGCTCTTTA 19  
||||| ||||| ||||| |||||  
Db 754 GGCATTGAGCGCTCTATA 772

RESULT 16  
US-10-282-122A-22599  
; Sequence 22599, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22599  
; LENGTH: 1110  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
US-10-282-122A-22599

Query Match  
Best Local Similarity 89.5%; Pred. No. 4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCATTGAGCGCTCTTTA 19  
||||| ||||| ||||| |||||  
Db 754 GGCATTGAGCGCTCTATA 772

RESULT 17  
US-09-882-227-89  
; Sequence 89, Application US/09882227  
; Publication No. US20030158396A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold



```

; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Coomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/802,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24)...(1130)
; US-09-882-227-89

Query Match          79.0%; Score 15.8; DB 10; Length 1170;
Best Local Similarity 89.5%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCTTTGGAGCGCTCTTTA 19
   ||||| ||||| |||||
Db 777 GCCTTTGAAGCGCTCTATA 795

RESULT 18
US-10-282-122A-39670/c
; Sequence 39670, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```

```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39670
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Salmonella typhi
; US-10-282-122A-39670

Query Match          79.0%; Score 15.8; DB 13; Length 1581;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCCTTTGGAGCGCTCTTTAA 20
   ||||| ||||| ||||| |||||
Db 1298 GCCTTTGGCGCTCTTTAA 1280

RESULT 19
US-10-194-163-995
; Sequence 995, Application US/10194163
; Publication No. US20020172976A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce Carter
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYNUCLEOTIDES
; AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/194,163
; FILING DATE: 04-Nov-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Basu, Shantanu
; REGISTRATION NUMBER: 43,318
; REFERENCE/DOCKET NUMBER: 529282000101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5995
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 995
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...4400
; SEQUENCE DESCRIPTION: SEQ ID NO: 995
; US-10-194-163-995

Query Match          79.0%; Score 15.8; DB 13; Length 4400;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCTTTGGAGCGCTCTTTA 19

```

```
Db 306 GGATTGGTGGCGCTCTTTA 324
|||||
RESULT 20
US-09-070-927A-146/c
; Sequence 146, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 146:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 146:
US-09-070-927A-146
Query Match 79.0%; Score 15.8; DB 9; Length 8546;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCGTTGGAGCGCTCTTTAA 20
|||||
Db 4375 GATTGGAGCGCTTTTAA 4357
|||||
RESULT 21
US-10-401-194-1/c
; Sequence 1, Application US/10401194
; Publication No. US20030219810A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Barnes, Glenn T.
; APPLICANT: Bertin, John
; TITLE OF INVENTION: POLYMORPHISMS IN THE HUMAN CARD4 GENE
; FILE REFERENCE: MPI02-041PRNM
; CURRENT APPLICATION NUMBER: US/10/401,194
```

```
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/368,184
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 68571
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-401-194-1
Query Match 79.0%; Score 15.8; DB 16; Length 68571;
Best Local Similarity 89.5%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGTTTGGAGCGCTCTTTA 19
|||||
Db 34348 GCGTTTGGAGCAGCTCTTTA 34330
|||||
RESULT 22
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKURO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 79.0%; Score 15.8; DB 9; Length 3309400;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGTTTGGAGCGCTCTTTA 19
|||||
Db 1404413 GCGTTTGTAGTGCCTTTA 1404395
|||||
RESULT 23
US-10-335-977-3971/c
; Sequence 3971, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1272.82 seconds  
(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819b-38

Perfect score: 20

Sequence: 1 ggctttggagcgctcttttaa 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	90.0	426	9	AI553971
C 2	18	90.0	516	9	AI910868
C 3	18	90.0	545	10	AW188109
C 4	18	90.0	891	12	BG717067
					-----
					AI553971 te49a10.x
					AI910868 wd20f03.x
					AW188109 xj92d12.x
					BG717067 60268914-

C	5	17.4	87.0	597	29	BX223249
C	6	17.4	87.0	615	14	CA354389
C	7	17	85.0	631	28	AQ756548
C	8	17	85.0	685	14	CF835325
C	9	16.8	84.0	277	10	BS29596
C	10	16.8	84.0	331	10	BE022711
C	11	16.8	84.0	366	9	AA569421
C	12	16.8	84.0	423	12	BM092062
C	13	16.8	84.0	425	12	BG511461
C	14	16.8	84.0	427	9	AI228821
C	15	16.8	84.0	497	14	CF863082
C	16	16.8	84.0	498	14	CA707288
C	17	16.8	84.0	513	13	BQ299456
C	18	16.8	84.0	516	10	BE115037
C	19	16.8	84.0	536	12	BI301021
C	20	16.8	84.0	542	12	EM953746
C	21	16.8	84.0	577	13	FU546888
C	22	16.8	84.0	586	9	AA603956
C	23	16.8	84.0	609	14	CD891133
C	24	16.8	84.0	636	10	BB480806
C	25	16.8	84.0	652	12	BJ624920
C	26	16.8	84.0	654	12	EM107891
C	27	16.8	84.0	663	9	AA683466
C	28	16.8	84.0	663	29	CE237499
C	29	16.8	84.0	671	12	BG839743
C	30	16.8	84.0	696	13	BY712910
C	31	16.8	84.0	704	14	CD898318
C	32	16.8	84.0	723	29	AG185598
C	33	16.8	84.0	798	29	CC906028
C	34	16.8	84.0	862	14	CB290368
C	35	16.8	84.0	870	29	CG954094
C	36	16.8	84.0	932	10	BE906019
C	37	16.8	84.0	934	10	BF103354
C	38	16.8	84.0	973	28	AQ743123
C	39	16.8	84.0	1004	11	AK013846
C	40	16.8	84.0	1268	11	AK084835
C	41	16.8	84.0	1277	11	AK089953
C	42	16.4	82.0	488	14	CF263777
C	43	16.4	82.0	538	28	AZ164174
C	44	16.4	82.0	553	28	AQ725857
C	45	16.4	82.0	557	14	CB213545
C	46	16.4	82.0	576	12	BJ003949
C	47	16.4	82.0	603	9	AV858696
C	48	16.4	82.0	611	9	AV967795
C	49	16.4	82.0	651	28	BZ202298
C	50	16.4	82.0	663	13	EM073951
C	51	16.4	82.0	693	29	CE632455
C	52	16.4	82.0	726	28	AZ330106
C	53	16.4	82.0	839	28	BZ560447
C	54	16.4	82.0	868	14	CB994896
C	55	16.4	82.0	899	29	CNS06VJA
C	56	16.4	82.0	1036	29	CG872551
C	57	16.4	82.0	1142	29	CNS02QHA
C	58	16	80.0	265	10	BF710761
C	59	16	80.0	527	28	CC130113
C	60	16	80.0	570	9	AI730174
C	61	16	80.0	1002	12	BG429601
C	62	15.8	79.0	160	29	CE705231
C	63	15.8	79.0	247	10	BB568268
C	64	15.8	79.0	261	28	AQ001200
C	65	15.8	79.0	276	28	AZ725377
C	66	15.8	79.0	294	28	AQ383346
C	67	15.8	79.0	343	28	CC40381
C	68	15.8	79.0	346	12	BM259540
C	69	15.8	79.0	368	10	BE453962
C	70	15.8	79.0	376	13	C27262
C	71	15.8	79.0	386	14	CF336070
C	72	15.8	79.0	389	9	AI769023
C	73	15.8	79.0	392	13	BY225009
C	74	15.8	79.0	397	29	CG633445
C	75	15.8	79.0	404	13	B0615908
C	76	15.8	79.0	412	12	BM500797
C	77	15.8	79.0	420	9	AJ475177

```

78 15.8 79.0 441 9 AL663926
79 15.8 79.0 444 14 N67764 za02f12.sl
80 15.8 79.0 450 12 B522369
81 15.8 79.0 453 9 AU006479
82 15.8 79.0 468 28 AZ402814
83 15.8 79.0 476 10 AW741895
84 15.8 79.0 480 9 AJ469281
85 15.8 79.0 486 28 BH476864
86 15.8 79.0 488 14 CB471125
87 15.8 79.0 491 13 BQ741564
88 15.8 79.0 492 10 AW654647
89 15.8 79.0 492 29 CC982363
90 15.8 79.0 494 12 BX073825
91 15.8 79.0 499 9 AI600575
92 15.8 79.0 506 14 D42733
93 15.8 79.0 528 28 AQ783813
94 15.8 79.0 535 13 BQ988978
95 15.8 79.0 536 29 CNS07GQ3
96 15.8 79.0 539 9 AA021775
97 15.8 79.0 541 12 BI542689
98 15.8 79.0 545 12 BM259212
99 15.8 79.0 545 13 BQ089196
100 15.8 79.0 548 10 B3475793
101 15.8 79.0 562 9 AV319296
102 15.8 79.0 595 13 BQ832203
103 15.8 79.0 600 29 AG212944
104 15.8 79.0 611 9 AV914262
105 15.8 79.0 612 28 B2202953
106 15.8 79.0 619 28 AY603720
107 15.8 79.0 625 29 CC862698
108 15.8 79.0 632 29 CG819239
109 15.8 79.0 639 12 B4855692
110 15.8 79.0 649 14 CA418535
111 15.8 79.0 651 10 BF435588
112 15.8 79.0 655 10 AW505948
113 15.8 79.0 658 28 AQ584860
114 15.8 79.0 669 13 BU207518
115 15.8 79.0 671 13 CA122050
116 15.8 79.0 676 28 CC431341
117 15.8 79.0 677 13 BU449519
118 15.8 79.0 678 9 AI296401
119 15.8 79.0 678 14 CA763938
120 15.8 79.0 686 13 BY762795
121 15.8 79.0 688 28 AZ326536
122 15.8 79.0 690 14 CA189995
123 15.8 79.0 695 13 BW042141
124 15.8 79.0 701 28 AQ006317
125 15.8 79.0 703 9 AI181066
126 15.8 79.0 703 10 BF500388
127 15.8 79.0 706 13 BU381297
128 15.8 79.0 712 12 BI358642
129 15.8 79.0 714 9 AW006171
130 15.8 79.0 723 28 BH976641
131 15.8 79.0 724 29 AG169417
132 15.8 79.0 751 13 CA122766
133 15.8 79.0 769 14 CF253965
134 15.8 79.0 772 13 BU243541
135 15.8 79.0 772 14 CK291961
136 15.8 79.0 800 28 CC096858
137 15.8 79.0 803 29 CG951694
138 15.8 79.0 808 29 CG092126
139 15.8 79.0 809 14 CD115756
140 15.8 79.0 817 29 CNS02XNO
141 15.8 79.0 863 29 CG177055
142 15.8 79.0 894 10 BF108273
143 15.8 79.0 937 13 BU281316
144 15.8 79.0 939 12 BI112382
145 15.8 79.0 957 29 CNS01HOM
146 15.8 79.0 966 29 CNS02C4C
147 15.8 79.0 972 29 CNS00F7M
148 15.8 79.0 973 29 CG115167
149 15.8 79.0 1038 28 AZ209312
150 15.8 79.0 1082 14 CK163088

AL663926 AL663926
N67764 za02f12.sl
B522369 B522369
AU006479 AU006479
AZ402814 1M0170J17
AW741895 uq97C09.Y
AJ469281 AJ469281
BH476864 BOHOI75TF
CB471125 sn3.A12
BQ741564 saq20B01.
AW654647 104664.MA
CC982363 ZUAD106TH
BX073825 MEAT76-D0
AI600575 486062D01
D42733 D42733 R1ce
AQ783813 HS_20C1_A
BQ988978 HF19F21.r
AL610077 Anopheles
AA021775 mh85e08.r
BI542689 949021C08
BM259212 952010H02
BQ089196 ko33a08.Y
BE475793 346048E11
AV919296 AV919296
BQ832203 LL6in2090
AG212944 Cryza sat
AV914262 AV914262
B2202953 CH230-304
AY603720 1M0423011
CG862698 NDL.64M14
CG819239 SOVEL04TH
BM485692 p5m1c.pk0
CA418535 UI-H-R21-
B343558 nac33e01
AW505948 GEA215.GI
AQ584860 RPCI-11-3
BU207518 604151929
CA122050 SCUFIR107
CG431341 PUBIK42TD
BU449519 603765641
AI296401 LP10272.5
CA763938 AF53-Rpf
BY762795 BY762795
AZ326536 1M0049A03
CA189995 SCCCLR1C0
BW042141 BW042141
AQ006317 CIT-HSP-2
AI181066 ud72h09.Y
BF500388 AT15261.5
BU381297 603860066
BI358642 GM28559.5
AW006171 wz93b05.x
BH976641 odf79h01.
AG169417 Pan trogl.
CA122766 SCULLR110
CF253965 mdv1110.b
BU243541 603779751
CK291961 EST754675
CC096858 CSU-X34.1
CG951694 MBEAQ16TR
CG092126 PUFQC19TD
CD115756 ME1-0038P
AL218589 Tetraodon
CG177055 PUKBX34TD
BU281316 603864148
BI112382 602500075
AI143831 Anopheles
AL190677 Tetraodon
ALC070240 Drosophila
CG115167 PUF7Q80TD
AZ209312 SP_0101A
CK163088 FGAS01570

```

## ALIGNMENTS

```

RESULT 1
AI553971/c
LOCUS      AI553971.1 426 bp mRNA linear EST 13-APR-1999
DEFINITION te49a10.x1 Soares NFL T GBC_S1 Homo sapiens cDNA clone
IMAGE:2090010.3, mRNA sequence.
ACCESSION  AI553971.1 GI:4486334
VERSION     AI553971.1
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (Bases 1 to 426)
AUTHORS     NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: ccaps-remail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 575 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source            1..426
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:2090010"
     /lab_host="DH10B"
     /clone_lib="Soares NFL T GBC_S1"
     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
     a modified polylinker; Site 1, Not 1; Site 2, Eco RI;
     Equal amounts of plasmid DNA from three normalized
     libraries (fetal lung NBH19W, testis NHT, and B-cell
     NCI-CCAP GCB1) were mixed, and ss circles were made in
     vitro. Following HAP purification, this DNA was used as

```

tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

## ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 426;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTT 18  
|||||  
Db 142 GCCTTTGGAGCGCTCTTT 125

## RESULT 2

AI910868/c  
LOCUS AI910868 516 bp mRNA linear EST 17-DEC-1999  
DEFINITION w22f03.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone  
IMAGE:2328701.3', mRNA sequence.

ACCESSION AI910868  
VERSION AI910868  
KEYWORDS EST.  
SOURCE AI910868.1 GI:5630604

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 516)

NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

UNPUBLISHED (1997)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 556 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 461.

Location/Qualifiers

1..516

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2328701"

/lab\_host="DH10B"

/clone\_lib="Soares NFL T GBC S1"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and 5-cell  
NCI CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo. "

## ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 515;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTT 18  
|||||  
Db 142 GCCTTTGGAGCGCTCTTT 125

## RESULT 3

AW188109/c

LOCUS AW188109 545 bp mRNA linear EST 22-NOV-1999

DEFINITION xj92di2.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone

IMAGE:2664695.3', mRNA sequence.

ACCESSION AW188109

VERSION AW188109

KEYWORDS EST.

SOURCE AW188109.1 GI:6462545

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 545)

NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

UNPUBLISHED (1997)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

This clone is available royalty-free through LNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 458.

Location/Qualifiers

1..545

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2664695"

/lab\_host="DH10B"

/clone\_lib="Soares NFL T GBC S1"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 545;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTT 18

|||||

Db 142 GCCTTTGGAGCGCTCTTT 125

## RESULT 4

BG717067

LOCUS BG717067 891 bp mRNA linear EST 08-MAY-2001

DEFINITION 602689141F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4821395.5',

mRNA sequence.

ACCESSION BG717067

VERSION BG717067.1 GI:13996254

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 891)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NECRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10728 row: a column: 12  
High quality sequence stop: 716.

FEATURES  
source  
1..891  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4821395"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC 97"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
5'-TTTTTTTTTTTTTTVN-3', size-selected for average  
insert size 2.2 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIMH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

## ORIGIN

Query Match 90.0%; Score 18; DB 12; Length 891;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTT 18  
|||||  
DB 408 GGCTTTGGAGCGCTCTT 425

RESULT 5  
BX223249/c

LOCUS  
DEFINITION  
Danio rerio genomic clone DKEX-50L13, genomic survey sequence.  
ACCESSION  
BX223249  
VERSION  
BX223249.1 GI:28055135  
KEYWORDS  
GSS:  
SOURCE  
Danio rerio (zebrafish)  
ORGANISM  
Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome  
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk) Unpublished  
This sequence was generated from the SP6 end of BAC 50L13. 50L13 is  
part of the Daniokey BAC Library created by R. Plasterk and N.V.  
Keygene. Further details:  
[http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).

FEATURES  
source

1..597  
Location/Qualifiers  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEX-50L13"  
/tissue\_type="Testis"  
/note="vector pIndigoBAC-536"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 29;  
Best Local Similarity 94.7%; Pred. No. 4.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTTA 19  
|||||  
DB 53 GGCTTTGGAGCTCTCTTA 35

RESULT 6  
CA354389  
LOCUS  
DEFINITION  
626102 NCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT77106\_C\_E03 5',  
mRNA sequence.  
ACCESSION  
CA354389  
VERSION  
CA354389.1 GI:24599576  
KEYWORDS  
EST.  
SOURCE  
Oncorhynchus mykiss (rainbow trout)  
ORGANISM  
Oncorhynchus mykiss

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 615)  
Rexroad, C.E. and Keele, J.W.

Sequence analysis of a rainbow trout normalized cDNA library  
Unpublished (2002)  
Contact: Rexroad CE  
USDA, ARS, National Center for Cool and Cold Water Aquaculture  
11876 Leetown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: [crexroad@nccwa.ars.usda.gov](mailto:crexroad@nccwa.ars.usda.gov)  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified by  
crossmatch v0.990329.

Seq primer: AGCGGATACATTTTCACACAGGA.

Location/Qualifiers  
1..615  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT77106\_C\_E03"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCWA 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from brain, gill, liver,  
spleen, muscle, and kidney."

FEATURES  
source

1..615  
Location/Qualifiers  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT77106\_C\_E03"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCWA 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from brain, gill, liver,  
spleen, muscle, and kidney."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 615;  
Best Local Similarity 94.7%; Pred. No. 4.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTTA 19  
|||||  
DB 101 GGCTTTGGAGCTCTCTTA 119

RESULT 7  
A0756548/c

LOCUS  
DEFINITION  
HS\_5373\_B2\_B03\_T7A RPII-11 Human Male BAC Library Homo sapiens  
genomic clone Plates=949 Col=6 Row=D, genomic survey sequence.  
ACCESSION  
A0756548  
VERSION  
A0756548.1 GI:5621206  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 631)  
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
10449764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For HAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

Plate: 949 row: D column: 6

Seq primer: T7

Class: BAC ends

High quality sequence stop: 631.

Location/Qualifiers

```
1. 631
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=949 Col=6 Row=D"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
(note="Vector: pBACe3.6; Site1: EcoRI; Site2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites")
```

## ORIGIN

Query Match 85.0%; Score 17; DB 28; Length 631;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CTTTGGAGCGCTCTTTA 19

Db 564 CTTTGGAGCGCTCTTTA 548

## RESULT 8

CF835325  
LOCUS  
DEFINITION  
UCRCS03\_01A20\_r Washington Navel Orange Shoot Meristem cDNA Library  
Citrus sinensis cDNA clone CS\_PEA01A20, mRNA sequence.

CF835325

CF835325.1 GI:38050875

EST.

SOURCE

Citrus sinensis

Citrus sinensis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Sapindales; Rutaceae; Citrus.

1 (bases 1 to 685)

Close,T.J., Roose,M.L., Federici,C.F., Mu,L., Penton,R.D.,

Wanamaker,S., Kim,H.R., Kudrna,D., Wing,R. and Yu.Y.

Development of EST Resources and New Genetic Markers for California

Citrus - Washington Navel Orange Shoot Meristem

Unpublished (2003)

Contact: Timothy Close

Department of Botany & Plant Sciences, University of California

Riverside, CA, 92521-0124

Tel: 9097873318

Fax: 9097874437

Email: timothy.close@ucr.edu

Seq primer: T3.

Location/Qualifiers

1. 685

/organism="Citrus sinensis"

## FEATURES

source

```
/mol_type="mRNA"
/cultivar="Parent Washington Navel"
/db_xref="taxon:2711"
/clone="CS_PEA01A20"
/tissue_type="Shoot meristem"
/dev_stage="10 Year old trees"
/lab_host="E. coli TJC121"
/clone_lib="Washington Navel Orange Shoot Meristem cDNA
Library"
```

(note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
Site 1: EcoRI; Site 2: XhoI; Parent Washington Navel  
Orange trees on Troyer rootstock (UCR 16K) were the source  
of tissue. Trees at UC Riverside Agricultural Operations,  
were planted October 12, 1992. In each of 17 reps one tree  
on Troyer rootstock was initially treated with Enzone, one  
with Alliette and Nemacure, and one was left untreated.  
These treatments were discontinued in 1998. At the time of  
sampling, there were differences in the apparent health  
and size of the trees on Troyer rootstock. Fall-flush  
shoots were sampled in early November 2002 to minimize the  
number of floral shoot meristems. Federici and Mu (Roose  
lab) harvested meristems only from trees that appeared to  
be healthy and had a large number of young shoot tips on  
the day of collection. The average weight of a meristem  
was about 2 mg. Federici noted that there were quite a few  
insects and signs of insect damage to the shoot tips.  
Mealy bugs, thrips and aphids were observed, plus a few  
very tiny fast moving insects that may have been mites or  
crawler stage of scale (although Federici did not see any  
mature scale). It was not difficult to avoid collecting  
most of these because they were easy to see with the  
dissecting microscope. It was harder to exclude the frass.  
Some frass was definitely retained in the samples. Tissues  
were snap frozen and then stored at -80C until further  
processing. Penton (Close lab) purified RNA by the phenol  
method described in J. Japanese Soc. Hort. Sci. 1996. 64  
(4): 809-814, purified poly(A) mRNA using a PolyAtrack  
mRNA Isolation System IV (Promega), produced a primary  
cDNA library using a lambda ZAP XR cDNA Synthesis Kit  
(Stratagene), then mass-excised one million pfu from the  
primary library to produce a phagemid population.  
Phagemids were plated, plasmid DNA purified, cDNA clones  
archived, and DNA sequences determined bi-directionally  
using an ABI3730 at the Arizona Genomics Institute,  
University of Arizona (Kim, Kudrna, Wing, Yu).  
Chromatogram files were downloaded by FTP to UC Riverside  
(by Close), then processed at UC Riverside (by Wanamaker,  
Close lab) using the HarvEST pipeline  
(http://harvest.ucr.edu) to remove vector and cloning  
oligo sequences and various contaminants, and to trim to a  
high quality region. Sequences that retained a phred 17  
region of at least 100 bases were deposited to GenBank."

## ORIGIN

Query Match 85.0%; Score 17; DB 14; Length 685;  
Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CTTTGGAGCGCTCTTTA 19

Db 297 CTTTGGAGCGCTCTTTA 313

## RESULT 9

BB529596

LOCUS

DEFINITION

BB529596 RIKEN full-length enriched, 0 day neonate lung Mus

musculus cDNA clone B33003D02.3', mRNA sequence.

ACCESSION

BB529596

VERSION

BB529596.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus





Db 268 GGCTTTGGAAAGCTCTTTAA 287

RESULT 11  
LOCUS  
DEFINITION  
AA569421 366 bp mRNA linear EST 25-AUG-1997  
TC-EST-037 Toxocara canis infective larva cDNA library Toxocara  
canis cDNA 5' similar to TC-aat-1; ADE/ATP translocase, mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AA569421  
AA569421.1 GI:2343311  
Toxocara canis  
Toxocara canis  
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida;  
Ascaridoidea; Toxocaridae; Toxocara.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 366)  
Tetteh,K.K.A., Loukas,A.C. and Maizels,R.M.  
Identification of Numerous Novel Genes Expressed by Infective  
Larvae of the Nematode Toxocara canis by Expressed Sequence Tag  
Analysis  
Unpublished (1997)  
Contact: Maizels RM  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, West Mains Road, Edinburgh EH93JT, Scotland  
Tel: +44 131 650 5511  
Fax: +44 131 650 5450  
Email: r.maizels@ed.ac.uk  
3' end of gene, EST starts approximately 250 aa downstream of start  
Insert Length: 380 Std Error: 0.00  
Seq primer: M13 Reverse.

FEATURES  
source  
1..366  
/organism="Toxocara canis"  
/mol\_type="mRNA"  
/db\_xref="taxon:6265"  
/clone\_lib="Toxocara canis infective larva cDNA library"  
/note="The cDNA library was constructed by Cindy Tripp  
(Heska Corporation, 1825 Sharp Point Drive, Fort Collins,  
Colorado 80525, USA) From 200,000 larval stage T. canis,  
using a single step guanidine-phenol-chloroform  
extraction, 265 ug total RNA was recovered, from which 6  
ug poly-A+ RNA was isolated by oligo-dT chromatography.  
cDNA synthesised from this mRNA was unidirectionally  
cloned into the Uni-Zap XR phage vector, using packaging  
extracts from Stratagene. The amplified library contained  
1.9 x 10<sup>9</sup> phage/ml with 91% recombinants."

ORIGIN  
Query Match 84.0%; Score 16.8; DB 9; Length 366;  
Best Local Similarity 90.0%; Pred.No. 7.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 78 GGCCTTGGAGCGCTCTTCAA 97

RESULT 12  
LOCUS  
DEFINITION  
BM092062 423 bp mRNA linear EST 29-NOV-2001  
sah07b01.y1 Gm-cl086 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl086-458 5' similar to TR:Q9SGS0 Q9SGS0 T23B18.6.; mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BM092062  
BM092062.1 GI:17021028  
Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 423)  
Shoemaker,R., Keim,P., Vodkin,L., Rypelding,J., Coryell,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McCamn,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Putative full length read  
vector to vector length is 442 This clone is available through:  
ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL  
35801 For further information call: (800)-533-4363 or contact via  
email: ccu@resgen.com.

FEATURES  
source  
1..423  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl086-458"  
/tissue type="young seeds (Williams 82)"  
/lab host="DH10B"  
/clone\_lib="Gm-cl086"  
/note="Vector: pBluescript II SK+; Site1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from very young seeds (less than 20mgs). The library was  
prepared using the Stratagene pBluescript II SK (+)  
library construction kit. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with an Xho I restriction site. Eco RI  
adaptors were ligated to the blunt-ended cDNA fragments  
followed by Xho I digestion. The cDNA insert is protected  
from Xho I digestion via methylation during first strand  
cDNA synthesis. The cDNA fragments were directionally  
cloned into the Eco RI-Xho I restriction site of the  
pBluescript vector. The ligated cDNA fragments were  
transformed into E.coli ElectroMax DH10B host cell. The  
library was constructed by Anu Khanna (Lila Vodkin lab,  
University of Illinois)."

ORIGIN  
Query Match 84.0%; Score 16.8; DB 12; Length 423;  
Best Local Similarity 90.0%; Pred.No. 7.7e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGCTTTGGAGCGCTCTTTAA 20  
|||||  
Db 213 GGCTTTGGAAAGCTCTTTAA 232

RESULT 13  
LOCUS  
DEFINITION  
BG511461 425 bp mRNA linear EST 28-NOV-2001  
sad02c02.y1 Gm-cl073 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl073-1275 5' similar to TR:Q9SGS0 Q9SGS0 T23E18.6.; mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BG511461  
BG511461.1 GI:13482118  
Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

```

REFERENCE
AUTHORS
  Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corvett,V.,
  Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
  Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
  Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
  Schurk,R., Ritter,B., Kohn,S., Stein,T., Jackson,Y., Cardenas,M.,
  McCann,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: ccu@resgen.com
  High quality sequence stop: 416.

FEATURES
  source
  1..425
    /organism="Glycine max"
    /mol_type="mRNA"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl073-1275"
    /tissue_type="seedlings induced for symptoms of SDS
    (Sudden Death Syndrome) disease"
    /dev_stage="2-3 weeks old"
    /lab_host="DH10B"
    /clone_lib="Gm-cl073"
    /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
    XhoI; The cDNA library was constructed from mRNA isolated
    from 2-3 week old seedlings that were induced for symptoms
    of SDS (Sudden Death Syndrome) disease by the
    translocation of culture filtrate of Fusarium solani f.
    sp. glycinis (Plant Cell Report 18:375-380). Cultivar
    Williams 82 is susceptible to the disease SDS. Plant
    tissue (expanded leaves, folded leaves, and new shoots)
    were collected at 1, 6, 24, and 48 hrs. after inoculation
    and their mRNA pooled equally for cDNA construction. The
    library was prepared using the Stratagene pBluescript II
    SK(+) library construction kit. Complementary DNA was
    synthesized from mRNA using a primer consisting of a
    poly(dT) sequence with an XhoI restriction site. EcoRI
    adaptors were ligated to the blunt-ended cDNA fragments
    followed by XhoI digestion. The cDNA insert is protected
    from XhoI digestion via methylation during first strand
    synthesis. The cDNA fragments were directionally cloned
    into the EcoRI-XhoI restriction site of the pBluescript
    vector. The ligated cDNA fragments were transformed into
    E.coli Electromax DH10B host cells. Plants were inoculated
    by Shuxian Li (Glen Hartman lab, University of Illinois).
    Library was constructed by Reena Philip and Steve Clough
    (Jila Vodkin lab, University of Illinois)."

ORIGIN
  Query Match      84.0%; Score 16.8; DB 12; Length 425;
  Best Local Similarity 90.0%; Pred. No. 7.7e-02;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCCTTGGAGCGCTCTTTAA 20
    |||||
Db 294 GGCCTTGGAGCGCTCTTTAA 313

RESULT 14
AI228821/c
LOCUS
  EST225516 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
  DEFINITION
  RBRCZ93 3' end, mRNA sequence.
  ACCESSION
  AI228821
  VERSION
  AI228821.1 GI:4135385

REFERENCE
  AUTHORS
  Tyler,B. Not Published
  TITLE
  Unpublished (2003)
  JOURNAL
  Contact: Tyler B
  COMMENT
  Tyler lab
  VBI
  1880 Pratt Dr., Blacksburg, VA 24061, USA
  Tel: 540-231-7318
  Email: bmttyler@vt.edu
  PCR Primers
  FORWARD: BK reverse primer
  BACKWARD: BK reverse primer
  Plate: 005 row: N column: 13
  Seq primer: BK reverse primer
  High quality sequence stop: 497.

FEATURES
  source
  1..497
    /organism="Phytophthora sojae"
    /mol_type="mRNA"

```

```

KEYWORDS
SOURCE
  Rattus sp.
  Rattus sp.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 427)
  Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
  Kerlavage,A.R. and Adams,M.D.
  Rat Genome Project: Generation of a Rat EST (REBT) Catalog & Rat
  Gene Index
  Unpublished (1998)
  On Oct 30, 1998 this sequence version replaced gi:3812708.
  Contact: Lee, NH
  The Institute for Genomic Research
  9712, Medical Center Drive, Rockville, MD 20850, USA
  Tel: (301)-838-3529
  Fax: (301)-838-0208
  Email: nhlee@tigr.org
  Seq primer: M13-21
  Location/Qualifiers
  1..427
    /organism="Rattus sp."
    /mol_type="mRNA"
    /db_xref="taxon:10118"
    /clone="RBRCZ93"
    /clone_lib="Normalized rat brain, Bento Soares"
    /note="Organ: brain; Vector: pRTT3pac; Site 1: EcoRI;
    Site 2: NotI"

ORIGIN
  Query Match      84.0%; Score 16.8; DB 9; Length 427;
  Best Local Similarity 90.0%; Pred. No. 7.7e-02;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCCTTGGAGCGCTCTTTAA 20
    |||||
Db 201 GGCCTTGGAGCGCTCTTTGA 182

RESULT 15
CF863082/c
LOCUS
  CF863082 497 bp mRNA linear EST 31-OCT-2003
  DEFINITION
  psZS005xN13f USDA-IPAPs:Expression of phytophthora sojae genes
  during infection and propagation_szs Phytophthora sojae cDNA clone
  szS005N13 5, mRNA sequence.
  ACCESSION
  CF863082
  VERSION
  CF863082.1 GI:38117708
  KEYWORDS
  EST.
  SOURCE
  Phytophthora sojae
  Phytophthora sojae
  Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
  Phytophthora.
  1 (bases 1 to 497)
  Tyler,B.
  Tyler,B. Not Published
  TITLE
  Unpublished (2003)
  JOURNAL
  Contact: Tyler B
  COMMENT
  Tyler lab
  VBI
  1880 Pratt Dr., Blacksburg, VA 24061, USA
  Tel: 540-231-7318
  Email: bmttyler@vt.edu
  PCR Primers
  FORWARD: BK reverse primer
  BACKWARD: BK reverse primer
  Plate: 005 row: N column: 13
  Seq primer: BK reverse primer
  High quality sequence stop: 497.

FEATURES
  source
  1..497
    /organism="Phytophthora sojae"
    /mol_type="mRNA"

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 604.235 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819b-39

Perfect score: 20

Sequence: 1 cagacgcacatagaccacacagg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 694C544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 180 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_pa.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_rts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_rts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_inv.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rdt.\*

36: em\_htg\_rnam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	189	9	MMDPAIF	Z32411 M.mullatta (
2	20	100.0	192	9	HUMMHCIITF	L31624 Human MHC C
3	20	100.0	222	9	AF074848	AF074848 Homo sapi
4	20	100.0	244	9	AF098794	AF098794 Homo sapi
5	20	100.0	246	9	A346471	AF346471 Homo sapi
6	20	100.0	246	9	HSOL03X2	X82390 H.sapiens M
7	20	100.0	246	9	HSLADPAL	X78198 H.sapiens H
8	20	100.0	249	9	HSDPAIRK	X96984 Homo sapien
9	20	100.0	267	6	AX237167	AX237167 Sequence
10	20	100.0	267	6	AX237352	AX237352 Sequence
11	20	100.0	268	9	HSLADPAX	X83610 H.sapiens H
12	20	100.0	272	6	AX237066	AX237066 Sequence
13	20	100.0	279	9	AF076284	AF076284 Homo sapi
14	20	100.0	279	9	AF076285	AF076285 Homo sapi
15	20	100.0	286	9	HSU87556	U87556 Homo sapien
16	20	100.0	287	9	AF015295	AF015295 Homo sapi
17	20	100.0	294	6	AX237304	AX237304 Sequence
18	20	100.0	294	6	AX237554	AX237554 Sequence
19	20	100.0	466	6	BD058339	BD058339 Secreted
20	20	100.0	476	6	AX884252	AX884252 Sequence
21	20	100.0	476	6	BD023862	BD023862 Sequence
22	20	100.0	576	6	AX884251	AX884251 Sequence
23	20	100.0	576	6	BD023861	BD023861 Sequence
24	20	100.0	661	6	E00485	E00485 DNA sequenc
25	20	100.0	661	6	I03086	I03086 Sequence 5
26	20	100.0	690	9	S40633	S40633 HLA Class I
27	20	100.0	818	9	HUMHDPXA	M27487 Homo sapien
28	20	100.0	1048	9	HSSBA1	X00457 Human mRNA
29	20	100.0	1140	6	I03088	I03088 Sequence 7
30	20	100.0	1201	6	E00484	E00484 DNA sequenc
31	20	100.0	1201	6	I03006	I03006 Sequence 6
32	20	100.0	1201	6	I03423	I03423 Sequence 6
33	20	100.0	1259	6	AX552229	AX552229 Sequence
34	20	100.0	2986	9	HUMHDPG02	M23904 Human MHC c
35	20	100.0	14646	9	HSLASBA	X03100 Human HLA-S
36	20	100.0	64380	9	EX120009	EX120009 Human DNA
37	20	100.0	106728	9	AL805913	AL805913 Human DNA
38	20	100.0	124899	9	AL645931	AL645931 Human DNA
39	20	100.0	181228	2	AC011086	AC011086 Homo sapi
40	20	100.0	187964	9	AL662824	AL662824 Human DNA
41	19	95.0	1480	6	AX780118	AX780118 Sequence
42	19	95.0	1480	6	AX780119	AX780119 Sequence
43	19	95.0	3888	10	AY082609	AY082609 Rattus no
44	19	95.0	4233	6	AR123273	AR123273 Sequence
45	19	95.0	4254	6	AX401748	AX401748 Sequence
46	19	95.0	4254	6	AX827531	AX827531 Sequence
47	19	95.0	4254	10	RATMDRM	M81855 Rat mdr mrn
48	19	95.0	168694	2	AC133679	AC133679 Rattus no
49	19	95.0	236777	2	AC094362	AC094362 Rattus no
50	18.4	92.0	189	9	AF026693	AF026693 Pan trogl
51	18.4	92.0	189	9	AF026694	AF026694 Pan trogl
52	18.4	92.0	189	9	AF026695	AF026695 Pongo pyg
53	18.4	92.0	189	9	AF026696	AF026696 Pongo pyg
54	18.4	92.0	189	9	AF026697	AF026697 Pongo pyg
55	18.4	92.0	189	9	AF026700	AF026700 Saimiri s
56	18.4	92.0	189	9	AF026701	AF026701 Gorilla g
57	18.4	92.0	189	9	AF026702	AF026702 Gorilla g
58	18.4	92.0	189	9	AF026703	AF026703 Macaca ar
59	18.4	92.0	189	9	AF026704	AF026704 Macaca fa
60	18.4	92.0	189	9	AF026705	AF026705 Macaca mu
61	18.4	92.0	189	9	AF026706	AF026706 Papio nam
62	18.4	92.0	189	9	AF026707	AF026707 Pan trogl
63	18.4	92.0	189	9	AF523200	AF523200 Aotus nan
64	18.4	92.0	192	9	HUMMHCIIAZ	L11643 Human MHC c
65	18.4	92.0	198	9	HUMMH7526X	M83909 Human MHC c

66	18.4	92.0	198	9	HUMHAMA1X	M83908 Human MHC C	139	17.4	87.0	147359	9	AP002345	Homo sapi
67	18.4	92.0	198	9	HUMHCB6BX	M83906 Human MHC C	C 140	17.4	87.0	148005	2	AC024944	Homo sapi
68	18.4	92.0	198	9	HUMHLK04X	M83907 Human MHC C	C 141	17.4	87.0	148193	9	AC027347	Homo sapi
69	18.4	92.0	214	9	HUMHC11AX	L11641 Human MHC C	C 142	17.4	87.0	148295	9	AC090510	Homo sapi
70	18.4	92.0	214	9	HUMHC11AX	L11642 Human MHC C	C 143	17.4	87.0	148476	9	AC130313	Homo sapi
71	18.4	92.0	226	9	HSDDPA126	X79478 H. sapiens M	C 144	17.4	87.0	149102	9	AC108095	Homo sapi
72	18.4	92.0	226	9	HSDDPA16B	X79475 H. sapiens M	C 145	17.4	87.0	149416	2	AC021985	Homo sapi
73	18.4	92.0	226	9	HSDDPA1AI	X79477 H. sapiens M	C 146	17.4	87.0	149731	9	HS162013	Homo sapi
74	18.4	92.0	226	9	HSDDPA1T3	X82394 H. sapiens M	C 147	17.4	87.0	149986	9	AC069256	Human DNA
75	18.4	92.0	246	9	HS0201X2	Z48473 H. sapiens M	C 148	17.4	87.0	150339	9	AC099515	Homo sapi
76	18.4	92.0	252	9	HSZ48473	AF165160 Homo sapi	C 149	17.4	87.0	150401	9	AC091907	Homo sapi
77	18.4	92.0	257	9	AF165160	AF118120 Homo sapi	C 150	17.4	87.0	151348	9	AC015807	Homo sapi
78	18.4	92.0	258	9	AF118120	AF092049 Homo sapi	C 151	17.4	87.0	152449	2	AC080175	Homo sapi
79	18.4	92.0	265	9	AF092049	AF013767 Homo sapi	C 152	17.4	87.0	153089	9	AC146510	Homo sapi
80	18.4	92.0	326	9	AF013767	BD058253 Secreterd	C 153	17.4	87.0	153291	2	AC117653	Homo sapi
81	18.4	92.0	396	6	BD058253	BC009956 Homo sapi	C 154	17.4	87.0	153496	2	AC087044	Homo sapi
82	18.4	92.0	1201	9	BC009956	AK126094 Homo sapi	C 155	17.4	87.0	153938	2	AC080050	Homo sapi
83	18.4	92.0	3161	9	AK126094	254246 Human DNA s	C 156	17.4	87.0	154387	2	AC018702	Homo sapi
84	18.4	92.0	36171	9	AK126094	269837 Human DNA s	C 157	17.4	87.0	154530	2	AC013596	Homo sapi
85	18.4	92.0	39756	9	HS1219B8	AC093404 Pan trogl	C 158	17.4	87.0	155307	2	CNS01DRG	Human chr
86	18.4	92.0	141287	2	AC093404	AL645949 Human DNA	C 159	17.4	87.0	156823	2	AC022051	Homo sapi
87	18.4	92.0	145151	9	AL645949	AL390053 Human DNA	C 160	17.4	87.0	158059	2	AC079947	Homo sapi
88	18.4	92.0	184590	9	AL390053	AC005516 Homo sapi	C 161	17.4	87.0	158813	9	AC080091	Homo sapi
89	18.4	92.0	200348	9	AC005516	AC108077 Mus muscu	C 162	17.4	87.0	159000	2	AC018427	Homo sapi
90	18.4	92.0	235097	2	AC108077	AF027333 Rattus no	C 163	17.4	87.0	159250	9	AC073628	Homo sapi
91	17.4	87.0	3056	10	AF027333	D38035 Rattus norv	C 164	17.4	87.0	159397	9	AC068544	Homo sapi
92	17.4	87.0	3511	10	RAT1T1	U19755 Mus domesti	C 165	17.4	87.0	159867	9	AL450303	Human DNA
93	17.4	87.0	11172	10	MMU19755	AL109617 Homo sapi	C 166	17.4	87.0	160421	2	AC025005	Homo sapi
94	17.4	87.0	36353	9	HS9P14	AL353808 Human DNA	C 167	17.4	87.0	160463	2	AC011566	Homo sapi
95	17.4	87.0	42622	9	AL353808	AL724232 Human DNA	C 168	17.4	87.0	160686	2	AC118002	Homo sapi
96	17.4	87.0	52664	9	AL724232	AC024919 Homo sapi	C 169	17.4	87.0	160762	9	AP000475	Homo sapi
97	17.4	87.0	62104	2	AC024919	AC090694 Homo sapi	C 170	17.4	87.0	160980	9	BS000075	Homo sapi
98	17.4	87.0	64744	2	AC090694	Continuation (4 of	C 171	17.4	87.0	161112	9	AC119800	Pan trogl
99	17.4	87.0	69471	2	AC117400	AL359759 Human DNA	C 172	17.4	87.0	162030	2	AC018601	Homo sapi
100	17.4	87.0	72978	9	AL359759	AC110056 Homo sapi	C 173	17.4	87.0	163640	9	AC093328	Homo sapi
101	17.4	87.0	74874	9	AL110056	AL022157 Human DNA	C 174	17.4	87.0	164020	2	AC021872	Homo sapi
102	17.4	87.0	80811	9	HS2323P24	AL645495 Human DNA	C 175	17.4	87.0	164173	2	AC016072	Homo sapi
103	17.4	87.0	81091	9	AL645495	AC139663 Homo sapi	C 176	17.4	87.0	164347	2	AC034124	Homo sapi
104	17.4	87.0	86563	9	AC139663	AC026865 Homo sapi	C 177	17.4	87.0	164479	2	AC103922	Homo sapi
105	17.4	87.0	88614	2	AC026865	AC095040 Homo sapi	C 178	17.4	87.0	164636	2	AC117783	Mus muscu
106	17.4	87.0	90171	9	AC095040	AC074289 Homo sapi	C 179	17.4	87.0	164652	2	AC026923	Homo sapi
107	17.4	87.0	92346	9	AC074289	EX005133 Human DNA	* 180	17.4	87.0	166302	2	AC121508	Mus muscu
108	17.4	87.0	96208	9	EX005133	AX695455 Sequence							
109	17.4	87.0	96596	6	AX695455	AL096770 Human DNA							
110	17.4	87.0	97392	9	HSB150A6	AL353706 Human DNA							
111	17.4	87.0	104679	9	AL353706	AL084013 Homo sapi							
112	17.4	87.0	105604	9	AC084013	AL662781 Homo sapi							
113	17.4	87.0	106642	9	AL662781	AC015923 Homo sapi							
114	17.4	87.0	109774	9	AC015923	Continuation (3 of							
115	17.4	87.0	110000	2	AC106723	AC012502 Homo sapi							
116	17.4	87.0	113799	9	AC012502	AC012355 Homo sapi							
117	17.4	87.0	120112	9	AC012355	AC092977 Homo sapi							
118	17.4	87.0	121158	9	AC092977	AL023875 Human DNA							
119	17.4	87.0	122325	9	HS1044017	AL357034 Human DNA							
120	17.4	87.0	127703	9	AL357034	AC091525 Homo sapi							
121	17.4	87.0	129740	9	AC091525	AL731777 Human DNA							
122	17.4	87.0	130110	9	AL731777	AC024911 Homo sapi							
123	17.4	87.0	130351	2	AC024911	Z82211 Human DNA s							
124	17.4	87.0	132229	9	HS447N6	AL356301 Homo sapi							
125	17.4	87.0	134584	2	AL356301	AL333485 Human chr							
126	17.4	87.0	135978	9	CNS01DV8	AC027567 Homo sapi							
127	17.4	87.0	137590	2	AC027567	AL157393 Human DNA							
128	17.4	87.0	138235	9	AL157393	AC021595 Homo sapi							
129	17.4	87.0	141854	2	AC021595	AC025257 Homo sapi							
130	17.4	87.0	142008	9	AC025257	AC008749 Homo sapi							
131	17.4	87.0	142439	9	AC008749	AC010082 Homo sapi							
132	17.4	87.0	143035	9	AC010082	AC087655 Homo sapi							
133	17.4	87.0	143244	2	AC087655	AC025157 Homo sapi							
134	17.4	87.0	143417	9	AC025157	AC105021 Homo sapi							
135	17.4	87.0	143552	9	AC105021	AC045666 Homo sapi							
136	17.4	87.0	145957	9	AC045666	AL358779 Human DNA							
137	17.4	87.0	147109	9	AL358779	AC025742 Homo sapi							
138	17.4	87.0	147156	2	AC025742								

## ALIGNMENTS

RESULT 1	MMDPALF	189 bp	DNA	linear	PRI 26-JUL-1995
MMDPALF	M.mulatta (clone Mamu-DPA1*0.01)	Mhc DP-alpha gene encoding major histocompatibility complex.			
LOCUS	Z32411.1	GI:4712620			
DEFINITION	major histocompatibility complex.				
ACCESSION	Z32411.1				
VERSION	232411.1				
KEYWORDS	Macaca mulatta (rhesus monkey)				
SOURCE	Macaca mulatta				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.				
REFERENCE	1 (bases 1 to 189)				
AUTHORS	Slijterdregt, B.L., Otting, N., Kenter, M. and Bontrop, R.E.				
TITLE	Allelic diversity at the Mhc-DP locus in rhesus macaques (Macaca mulatta)				
JOURNAL	Immunogenetics 41 (1), 29-37 (1995)				
MEDLINE	95104902				
PUBMED	7806271				
REFERENCE	2 (bases 1 to 189)				
AUTHORS	Slijterdregt, B.L.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-APR-1994) Bastiaan B.L. Slijterdregt, Immunobiology, Biomedical Primate, Research Center-TNO, Lange Kleiweg 151, Rijswijk, Zuid-Holland, 2288 GJ, The Netherlands				

```

FEATURES             source
  Location/Qualifiers
    1..189
      /organism="Macaca mulatta"
      /mol_type="genomic DNA"
      /isolate="2849"
      /db_xref="taxon:9544"
      /clone="Mamu-DPA1*0101"
      /cell_type="Lymphocyte"
      /tissue_type="Blood"
    1..189
      /gene="Mhc DP-alpha"
      /note="exon 2"
      /codon_start=1
      /product="major histocompatibility complex"
      /protein_id="CAA83527.1"
      /db_xref="GI:471261"
      /db_xref="GOA:Q30617"
      /db_xref="SPTREMBL:Q30617"
      /translation="FVQTHRTGTFMFEFDEMFYVLDLKKETVWHLSEFGQAFSFE
      AQGSLANIALNNNLITLI"

ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 7 CAGACGCATAGACCAACAGG 26

RESULT 2
HUMHCLITF
LOCUS Human MHC class II gene (DPA1*TF) linear PRI 03-MAY-1995
ACCESSION L31624
VERSION L31624
KEYWORDS cell surface glycoprotein; class II gene; integral membrane
SOURCE protein; major histocompatibility complex.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Meyer, C.G., May, J., Spaake, D. and Schnittger, L.
TITLE DPAl*02012: a DPAl*0201-related Mhc class II allele in west Africa
JOURNAL Immunogenetics 40 (4), 309 (1994)
MEDLINE 94364641
PUBMED 8082895
COMMENT Original source text: Homo sapiens male blood DNA.
FEATURES             source
  Location/Qualifiers
    1..192
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /sex="male"
      /cell_type="Lymphocyte"
      /tissue_type="blood"

ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 13 CAGACGCATAGACCAACAGG 32

RESULT 3
AF074848
LOCUS Homo sapiens MHC class II antigen (HLA-DPA1) gene, HLA-DPA1*01032
allele, partial cds.
ACCESSION AF074848
VERSION AF074848.1 GI:5106402
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Steiner, L.L., Wu, J., Noreen, H.J., Moehlenkamp, C., Cavalli, A.,
Davidson, M., Johnson, S., Winden, T., Segali, M., Begovich, A.B. and
Williams, T.M.
TITLE Four new DP alleles identified in a study of 500 unrelated bone
marrow donor-recipient pairs
JOURNAL Tissue Antigens 53 (2), 201-206 (1999)
MEDLINE 99189032
PUBMED 10090623
REFERENCE 2 (bases 1 to 222)
AUTHORS Steiner, L. and Begovich, A.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-1998) Human Genetics, Roche Molecular Systems,
Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA
FEATURES             source
  Location/Qualifiers
    1..222
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="6"
      /contig="222"
      /gene="HLA-DPA1"
      /allele="HLA-DPA1*01032"
      /number=2
    <1..>222
      /gene="HLA-DPA1"
      /product="MHC Class II antigen"
    <1..>222
      /gene="HLA-DPA1"
      /codon_start=1
      /product="MHC class II antigen"
      /protein_id="AAD39685.1"
      /db_xref="GI:5106403"
      /translation="AFVQTHRTGTFMFEFDEMFYVLDLKKETVWHLSEFGQAFSFE
      EAQGLANIALNNNLITLIQRSH*QATN"
    1..222
      /gene="HLA-DPA1"
      /number=2

ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 10 CAGACGCATAGACCAACAGG 29

RESULT 4
AF098794
LOCUS Homo sapiens MHC class II antigen (HLA-DPA1) gene, exon 2 and
partial cds.
ACCESSION AF098794
VERSION AF098794.1 GI:3859561
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
McTernan, C.L., Mijovic, C.H., Cockram, C.S. and Barnett, A.H.
TITLE The nucleotide sequence of two new DP alleles, DPA1*02015 and

```

```

DPB1*8401, identified in a Chinese subject
Tissue Antigens 56 (1), 95-98 (2000);
20412568
PUBMED
10958363
2 (bases 1 to 244)
Perry, C.L., Mijovic, C.E., Cockram, C.S. and Barnett, A.H.
AUTHORS
Direct Submission
TITLE
Submitted (14-OCT-1998) Medicine, Birmingham University, Clinical
JOURNAL
Research Block, Queen Elizabeth Hospital, Edgbaston, Birmingham B15
2TH, England
FEATURES
    source
        location/Qualifiers
            1..244
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21.1"
            <1..>244
            /gene="HLA-DPA1"
            <1..>244
            /gene="HLA-DPA1"
            /product="MHC class II antigen"
            <1..>244
            /gene="HLA-DPA1"
            /codon_start=1
            /product="MHC class II antigen"
            /protein_id="AAC72845.1"
            /db_xref="GI:3859562"
            /translation="DHVSTYAAAFVQTHRPTGEFMFEDDEQFYVDLDKKEVTWHLRE
            FGRAFSPEAQGLANIALNNLNTLIQRSHNTQAN"
            <1..244
            /gene="HLA-DPA1"
            /number=2
gene
    mRNA
        <1..>244
        /gene="HLA-DPA1"
        /product="MHC class II antigen"
        <1..>244
        /gene="HLA-DPA1"
        /codon_start=3
        /product="MHC class II antigen"
        /protein_id="AAK27152.1"
        /db_xref="GI:13448662"
        /translation="DHVSTYAAAFVQTHRPTGEFMFEDDEQFYVDLDKKEVTWHLRE
        FGRAFSPEAQGLANIALNNLNTLIQRSHNTQATN"
            1..246
            /gene="HLA-DPA1"
            /number=2
    exon
    CDS
        <1..>244
        /gene="HLA-DPA1"
        /product="MHC class II antigen"
        <1..>244
        /gene="HLA-DPA1"
        /codon_start=3
        /product="MHC class II antigen"
        /protein_id="AAK27152.1"
        /db_xref="GI:13448662"
        /translation="DHVSTYAAAFVQTHRPTGEFMFEDDEQFYVDLDKKEVTWHLRE
        FGRAFSPEAQGLANIALNNLNTLIQRSHNTQATN"
            1..246
            /gene="HLA-DPA1"
            /number=2
    ORIGIN
        Query Match 100.0%; Score 20; DB 9; Length 246;
        Best Local Similarity 100.0%; Pred. No. 1.3;
        Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CAGACGCATAGACCAACAGG 20
    ||||||||||||||||||
Db 33 CAGACGCATAGACCAACAGG 52
    ||||||||||||||||||
RESULT 6
LOCUS HS0103X2 246 bp DNA linear PRI 14-JUL-1995
DEFINITION H.sapiens MHC class II HLA-DPA1*0103 gene (exon 2).
ACCESSION X82390
VERSION X82390.1 GI:565028
KEYWORDS MHC class II HLA DPA1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Rozemuller, E.H., Bouwens, A.G., van Oort, E., Versluis, L.F.,
Marsh, S.G., Bodmer, J.G. and Tilius, M.G.
Sequencing-based typing reveals new insight in HLA-DPA1
polymorphism
Tissue Antigens 45 (1), 57-62 (1995)
JOURNAL MEDLINE 95242313
PUBMED 7725312
REFERENCE 2 (bases 1 to 246)
AUTHORS Rozemuller, E.H.
Direct Submission
TITLE Submitted (19-OCT-1994) E.H. Rozemuller, Diagnostic DNA Laboratory,
JOURNAL University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
Utrecht, NETHERLANDS
COMMENT Related sequences: S52453 and D14344.
FEATURES
    source
        location/Qualifiers
            1..246
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /plotype="DPA1*0103"
            /cell_line="LB and LG2"
            1..246
            /gene="MHC class II HLA-DPA1 gene"
            1..246
            /gene="MHC class II HLA-DPA1 gene"
            /number=2
    gene
    exon
    ORIGIN
        Query Match 100.0%; Score 20; DB 9; Length 246;
        Best Local Similarity 100.0%; Pred. No. 1.3;
        Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CAGACGCATAGACCAACAGG 20
    ||||||||||||||||||
Db 31 CAGACGCATAGACCAACAGG 50
    ||||||||||||||||||
RESULT 5
AF346471
LOCUS AF346471 246 bp DNA linear PRI 27-MAR-2001
DEFINITION Homo sapiens MHC class II antigen (HLA-DPA1) gene, HLA-DPA1-new
allele, partial cds.
ACCESSION AF346471
VERSION AF346471.1 GI:13448661
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Grams, S.E., Begovich, A. and Mangaccat, J.
TITLE One new DPA1 Allele
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 246)
AUTHORS Grams, S.E., Begovich, A. and Mangaccat, J.
Direct Submission
TITLE Submitted (07-FEB-2001) Human Genetics, Roche Molecular Systems,
JOURNAL 1145 Atlantic Ave., Alameda, CA 94501, USA
FEATURES
    source
        location/Qualifiers
            1..246
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p"
            <1..>246
gene

```

```

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 33 CAGACGCATAGACCAACAGG 52

RESULT 7
HSLADPAL          HSLADPAL          246 bp      DNA      linear      PRI 14-JUL-1995
LOCUS             H.sapiens HLA-DPAL gene, exon 2.
DEFINITION        X78198
ACCESSION         X78198.1 GI:461355
VERSION           DPAL*01new; HLA-DPAL gene.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1 (bases 1 to 246)
Rozenmuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,
Marsh,S.G., Bodner,J.G. and Tillanus,M.G.
Sequencing-based typing reveals new insight in HLA-DPAL
polymorphism
Tissue Antigens 45 (1), 57-62 (1995)
JOURNAL           95242313
MEDLINE           7725312
PUBMED            7725312
REFERENCE         2 (bases 1 to 246)
Rozenmuller,E.H.
Direct Submission
TITLE            Submitted (15-MAR-1994) E.H. Rozenmuller, Diagnostic DNA Lab,
JOURNAL           University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
Utrecht, NETHERLANDS
FEATURES          Location/Qualifiers
                    source
                    1..246
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /isolate="SK"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /haplotype="DPAL*01"
                        /gene="HLA-DPAL"
                        /product="DPAL*01new"
                        /number=2

gene             1..246
exon             1..246
                /gene="HLA-DPAL"
                /product="DPAL*01new"
                /number=2

ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 33 CAGACGCATAGACCAACAGG 52

RESULT 8
HSDPALRK          HSDPALRK          249 bp      DNA      linear      PRI 24-NOV-2000
LOCUS             Homo sapiens HLA-DPAL gene, exon 2, isolate DNA RX.
DEFINITION        X96984
ACCESSION         X96984.1 GI:1580762
VERSION           antigen; MHC; MHC class II; MHC class II DPAL gene.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1
May, J., Kreschmer, C., Schnitzger, L., Striecker, R., Kremoner, P.G.
and Meyer, C.G.
DPAL*0105, a novel DPAL variant in a negrois population
JOURNAL           Tissue Antigens 48, 693-694 (1996)
REFERENCE         2 (bases 1 to 249)
Meyer, C.G.
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (29-MAR-1996) C.G. Meyer, Insitute of Tropical Medicine
Berlin, Mol. Biol., Engelidamm 62, 10179 Berlin, FRG
FEATURES          Location/Qualifiers
                    source
                    1..249
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /isolate="DNA RK from Gabon"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /tissue_type="blood"
                        /dev stage="adult"
                        /number=2
                    10..84
                        /product="MHC-class II DPAL antigen"
                        /number=2
                    exon
                    Query Match      100.0%; Score 20; DB 9; Length 249;
                    Best Local Similarity 100.0%; Pred. No. 1.3;
                    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 13 CAGACGCATAGACCAACAGG 32

RESULT 9
AX237167          AX237167          267 bp      DNA      linear      PAT 26-SEP-2001
LOCUS             Sequence 143 from Patent WO0164886.
DEFINITION        AX237167
ACCESSION         AX237167
VERSION           AX237167.1 GI:15796721
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1
Gaiger, A., Algate, P.A. and Mannion, J.
Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
Patent: WO 0164886-A 143 07-SEP-2001;
CORIXA CORPORATION (US)
JOURNAL           Location/Qualifiers
FEATURES          source
                    1..267
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 10
AX237352          AX237352          267 bp      DNA      linear      PAT 26-SEP-2001
LOCUS             Sequence 328 from Patent WO0164886.
DEFINITION        AX237352
ACCESSION         AX237352
VERSION           AX237352.1 GI:15796906
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1
Gaiger, A., Algate, P.A. and Mannion, J.
Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
Patent: WO 0164886-A 143 07-SEP-2001;
CORIXA CORPORATION (US)
JOURNAL           Location/Qualifiers
FEATURES          source
                    1..267
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 10
AX237352          AX237352          267 bp      DNA      linear      PAT 26-SEP-2001
LOCUS             Sequence 328 from Patent WO0164886.
DEFINITION        AX237352
ACCESSION         AX237352
VERSION           AX237352.1 GI:15796906
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1
Gaiger, A., Algate, P.A. and Mannion, J.
Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
Patent: WO 0164886-A 143 07-SEP-2001;
CORIXA CORPORATION (US)
JOURNAL           Location/Qualifiers
FEATURES          source
                    1..267
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 33 CAGACGCATAGACCAACAGG 52

RESULT 8
HSDPALRK          HSDPALRK          249 bp      DNA      linear      PRI 24-NOV-2000
LOCUS             Homo sapiens HLA-DPAL gene, exon 2, isolate DNA RX.
DEFINITION        X96984
ACCESSION         X96984.1 GI:1580762
VERSION           antigen; MHC; MHC class II; MHC class II DPAL gene.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE            1
May, J., Kreschmer, C., Schnitzger, L., Striecker, R., Kremoner, P.G.
and Meyer, C.G.
DPAL*0105, a novel DPAL variant in a negrois population
JOURNAL           Tissue Antigens 48, 693-694 (1996)
REFERENCE         2 (bases 1 to 249)
Meyer, C.G.
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

of hematological malignancies
JOURNAL Patent: WO 0164886-A 328 07-SEP-2001;
FEATURES CORIXA CORPORATION (US)
SOURCE Location/Qualifiers
1..267
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
|||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 11
LOCUS HSHLADPAX 268 bp DNA linear PRI 14-SEP-1995
DEFINITION H.sapiens HLA-DPA1 gene.
ACCESSION X83610
VERSION X83610.1 GI:987073
KEYWORDS HLA-DPA1 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1.
Versluis, L.P., Verduyn, W., Van der Zwan, A., Oudshoorn, M. and
Tilanus, M.G.J.
An update of the exon 2 sequence of the HLA-DPA1*02012 allele
Tissue Antigens 46 (3 Pt 1), 206-207 (1995)
96097411
PUBMED 8525481
REFERENCE 2 (bases 1 to 268)
AUTHORS Tilanus, M.G.J.
TITLE Direct Submission
JOURNAL Submitted (20-DEC-1994) M.G.J. Tilanus, Diagnostic DNA Laboratory,
Academic Hospital Utrecht, Heidelberglaan 100, PO Box 85500, 3508
GA Utrecht, NETHERLANDS
Related sequence: L31524.
COMMENT Location/Qualifiers
1..268
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/haplotype="DPA1*02012"
<1..18
introns
19..264
/gene="HLA-DPA1"
265..268
exons
19..264
/gene="HLA-DPA1"
265..268
introns

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
|||||
Db 51 CAGACGCATAGACCAACAGG 70

RESULT 12
LOCUS AX237066 272 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 42 from Patent WO0164886.
ACCESSION AX237066

of hematological malignancies
JOURNAL Patent: WO 0164886-A 328 07-SEP-2001;
FEATURES CORIXA CORPORATION (US)
SOURCE Location/Qualifiers
1..272
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
|||||
Db 166 CAGACGCATAGACCAACAGG 185

RESULT 13
LOCUS AF076284 279 bp DNA linear PRI 07-JUL-1999
DEFINITION Homo sapiens isolate 913 MHC class II antigen (HLA-DPA1) gene,
partial cds.
ACCESSION AF076284
VERSION AF076284.1 GI:5381292
KEYWORDS
SOURCE Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 279)
Varney, M.D., Gavrilidis, A. and Abbott, W.
DPA1 Polymorphism in Polynesians
Unpublished
REFERENCE 2 (bases 1 to 279)
AUTHORS Varney, M.D., Gavrilidis, A. and Abbott, W.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) Tissue Typing, Royal Melbourne Hospital,
Grattan Street, Parkville, Vic 3050, Australia
Location/Qualifiers
1..279
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="913"
/db_xref="taxon:9606"
/chromosome="6"
<1..279
/gene="HLA-DPA1"
<1..279
/gene="HLA-DPA1"
/product="MHC class II antigen"
<1..279
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAD42927.1"
/db_xref="GI:5381293"
/translation="HVSITAAAFVQTHRPTGEFMFEDEMFYVDLDKKTVMHLREF
GQTFPEAQGLANIALNNNLTIQRSNHTQATNGTPTLCIFLCSTP"
1..279
/gene="HLA-DPA1"
/number=2

ORIGIN

```



```

Query Match      100.0%; Score 20; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 28 CAGACGCATAGACCAACAGG 47

RESULT 14
LOCUS AF076285 279 bp DNA linear PRI 07-JUL-1999
DEFINITION Homo sapiens isolate 63 MHC class II antigen (HLA-DPA1) gene,
partial cds.
ACCESSION AF076285
VERSION AF076285.1 GI:5381294
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 279)
AUTHORS Varney,M.D., Gavrilidis,A. and Abbott,W.
TITLE DPA1 Polymorphism in Polynesians
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 279)
AUTHORS Varney,M.D., Gavrilidis,A. and Abbott,W.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) Tissue Typing, Royal Melbourne Hospital,
Grattan Street, Parkville, Vic 3050, Australia
FEATURES
    source
        Location/Qualifiers
            1..279
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /isolate="63"
                /db_xref="taxon:9606"
                /chromosome="6"
                <1..>279
                /gene="HLA-DPA1"
                <1..>279
                /gene="HLA-DPA1"
                <1..>279
                /product="MHC class II antigen"
                <1..>279
                /codon_start=1
                /product="MHC class II antigen"
                /protein_id="AAD42928.1"
                /db_xref="GI:5381295"
                /translation="HVSIVYAAFVQTHRPTGFMFPEDEDEQFYVDLKKETVHLEEF
                GRAFSFEAQGGLANIAILNNLNTLIQRSNHTQAANGPYLCFLCISPT"
            1..279
                /gene="HLA-DPA1"
                /number=2

Query Match      100.0%; Score 20; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 28 CAGACGCATAGACCAACAGG 47

RESULT 15
LOCUS HSU87556 286 bp DNA linear PRI 20-JAN-1998
DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2,
partial cds.
ACCESSION U87556
VERSION U87556.1 GI:2760313
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 286)
AUTHORS Steiner,L., Begovich,A. and Suraj,V.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
COMMENT On Jan 8, 1998 this sequence version replaced gi:1842112.
FEATURES
    source
        Location/Qualifiers
            1..286
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="6"
                <1..>286
                /gene="HLA-D"
                <16..>261
                /gene="HLA-D"
                /note="MHC class II HLA-DPA1 antigen"
                /codon_start=3
                /protein_id="AAB97110.1"
                /db_xref="GI:2795772"
                /translation="DHVSTYAAFVQTHRPTGFMFPEDEDEQFYVDLKKETVHLEE
                FGGAPSEEAQGGLANIAILNNLNTLIQRSNHTQATN"
            16..261
                /gene="HLA-D"
                /number=2

Query Match      100.0%; Score 20; DB 9; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 48 CAGACGCATAGACCAACAGG 67

RESULT 16
LOCUS AF015295 287 bp DNA linear PRI 29-SEP-1998
DEFINITION Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*02013
allele), exon 2 and partial cds.
ACCESSION AF015295
VERSION AF015295.1 GI:3660653
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 287)
AUTHORS Steiner,L.L., Cavalli,A., Zimmerman,P.A., Boatin,B.A.,
Titanji,V.P., Bradley,J.E., Lucius,R., Nutman,T.B. and
Begovich,A.B.
TITLE Three new DP alleles identified in sub-Saharan Africa: DPB1*7401,
DPA1*02013, and DPA1*0302
JOURNAL Tissue Antigens 51 (6), 653-657 (1998)
MEDLINE 98357732
PUBMED 9694359
REFERENCE 2 (bases 1 to 287)
AUTHORS Steiner,L., Begovich,A. and Zimmerman,P.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1997) Human Genetics, Roche Molecular Systems,
Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA
FEATURES
    source
        Location/Qualifiers
            1..287
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="6"
                <1..>287

```

```

/genes="HLA-DPA1"
/allele="HLA-DPA1*02013"
<1..15
/genes="HLA-DPA1"
/number=1
<16..>261
/genes="HLA-DPA1"
/product="MHC class II antigen"
<16..>261
/genes="HLA-DPA1"
/codon_start=3
/product="MHC class II antigen"
/protein_id="AAC61669.1"
/db_xref="GI:3660654"
/translation="DRVSTYAAFVQTHRTGTFMFDEQFYVDLKKETVWHLEE
FGRAFSFEAQGLANLAINNNLTIQRSHNTQAAAN"
16..261
/genes="HLA-DPA1"
/number=2
262..>287
/genes="HLA-DPA1"
/number=2

Query Match 100.0%; Score 20; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 48 CAGACGCATAGACCAACAGG 67

RESULT 17
AX237304/c
LOCUS 294 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 280 from Patent WO0164886.
ACCESSION AX237304
VERSION AX237304.1 GI:15796858
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A, 280 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source
location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 107 CAGACGCATAGACCAACAGG 88

RESULT 18
AX237554/c
LOCUS 294 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 530 from Patent WO0164886.
ACCESSION AX237554
VERSION AX237554.1 GI:15797108
KEYWORDS

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A, 530 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source
location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 107 CAGACGCATAGACCAACAGG 88

RESULT 19
BD058339
LOCUS 466 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (SESTs).
ACCESSION BD058339
VERSION BD058339.1 GI:22603945
KEYWORDS Jp 2001519666-A/194.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Treacy, M., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,
Treacy, M., Spaulding, V. and Agostino, M.J.
TITLE Secreted expressed sequence tags (SESTs)
JOURNAL Patent: JP 2001519666-A, 194 23-OCT-2001;
GENETICS INSTITUTE INC
COMMENT FN CP 2001519666-A/194
PD 23-OCT-2001
PF 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
1..466
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"

FEATURES
source
location/Qualifiers
1..466
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 152 CAGACGCATAGACCAACAGG 171

RESULT 20
AX884252

```

```

source
1. .476
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
        |||||
DB      227 CAGACGCATAGACCAACAGG 246

RESULT 22
AX884251          AX884251          576 bp      DNA      linear      PAT 18-DEC-2003
LOCUS             Sequence 114 from Patent EP1033401.
DEFINITION        AX884251
ACCESSION         AX884251
VERSION           AX884251.1 GI:40039225
KEYWORDS
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE             Expressed sequence tags and encoded human proteins
JOURNAL           Patent : EP 1033401-A 114 06-SEP-2000;
GENSET            Genset (FR)

FEATURES
Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
195..>575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF98365.1"
/db_xref="GI:40039226"
/translation="MRPEDMFHRAVILRALSLAFLLSLRGAGAIKADHVSYAAFAV
QTRPTGEFMFFDEMFYDLDKKETVMHLTEFGQAFSEAQGLAIAILNNLIN
TLIQRSNHTQATNDPDEVTVFPKEP"
sig_peptide       195..287
/note="score 10.1 seq SLAFLLSLRGAGA/IX"

ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
        |||||
DB      327 CAGACGCATAGACCAACAGG 346

RESULT 23
BD023861          BD023861          576 bp      DNA      linear      PAT 27-AUG-2002
LOCUS             Sequence tag and encoded human protein.
DEFINITION        BD023861
ACCESSION         BD023861
VERSION           BD023861.1 GI:22565084
KEYWORDS          JP 2001269182-A/107.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 576)
AUTHORS           Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE             Sequence tag and encoded human protein
JOURNAL           Patent: JP 2001269182-A 107 02-OCT-2001;
GENSET            GENSTT

```

```

COMMENT OS Homo sapiens (human)
PN JP 200-269182-A/107
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PI 26-FEB-1999 US 60/122487
PI JEAN RAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC score 10.1
CC seq SLAFLSLRGAGA/IK
FH Key Location/Qualifiers
FT CDS 195..575
FT sig_peptide 195..287.
FEATURES
source
1..576
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 327 CAGACGCATAGACCAACAGG 346
|||||

RESULT 24
E00485
LOCUS 661 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of a fragment of pSBalpha-318.
ACCESSION E00485
VERSION E00485.1 GI:2168768
KEYWORDS JP 1985226888-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 661)
AUTHORS Edwaado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL Patent: JP 1985226888-A 3 12-NOV-1985;
CETUS CORP
COMMENT OS Human {Homo sapiens}
PN JP 1985226888-A/3
PD 12-NOV-1985
PF 20-MAR-1985 JP 1985054705
PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 603263 PI
EDWAADO RUISU SHIERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
KOOREI HAWAADO REBENSON, HENRII RAPOPORUTO
PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/50,G01N33/532,
PC G01N33/58
PC (C07D519/00,C07D493/04,C07D495/04),(C07D519/00,C07D493/04, PC
C07D493/10);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=Blood;
CC *source: cell_type=B cell;
CC *source: library=cDNA library;
FH Key Location/Qualifiers
FT misc_feature 1..661
FT note=a fragment derived from pSBalpha-318
FT Location/Qualifiers
1..661
/organism="unidentified"
FEATURES
source

```

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 211 CAGACGCATAGACCAACAGG 230  
 |||||

## RESULT 25

LOCUS 103086 661 bp ss-DNA linear PAT 21-MAY-1993  
 DEFINITION Sequence 5 from Patent US 4582789.  
 ACCESSION 103086  
 VERSION 103086.1 GI:268242  
 KEYWORDS  
 SOURCE Unknown.

ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 661)  
 AUTHORS Sheidos,E.I. III, Levenson,C.H., Mullis,K.B. and Rapoport,H.  
 TITLE Process for labeling nucleic acids using psoralen derivatives  
 JOURNAL Patent: US 4582789-A 5 15-APR-1986;  
 Cetus Corporation; Emeryville, CA

FEATURES  
 source  
 1..661  
 Location/Qualifiers  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 211 CAGACGCATAGACCAACAGG 230  
 |||||

## RESULT 26

LOCUS S40633 630 bp DNA linear PRI 06-MAY-1993  
 DEFINITION HLA class II: DPA1 (DPA1\*0101) [human, Genomic, 630 nt].  
 ACCESSION S40633  
 VERSION S40633.1 GI:1679890

KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 630)  
 AUTHORS Marsh,S.G. and Bodmer,J.G.  
 TITLE HLA class II nucleotide sequences, 1991  
 JOURNAL Immunogenetics 33 (5-6), 321-334 (1991)

MEDLINE 91267561  
 PUBMED 1904836  
 REMARK Genbank staff at the National Library of Medicine created this  
 entry [NCBI gibsseq 40633] from the original journal article.  
 This sequence comes from Figure 15.  
 On Nov 21, 1996 this sequence version replaced gi:1619630.  
 Region: HLA class II.

FEATURES  
 source  
 1..630  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1..630  
 /gene="DPA1"  
 /allele="DPA1\*0101"

## FEATURES

source

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||  
 DB 40 CAGACGCATAGACCAACAGG 59

## RESULT 27

LOCUS HUMHMDPA 818 bp mRNA linear PRI 07-MAR-1995  
 DEFINITION Homo sapiens MHC class II DPw3-alpha-1 chain mRNA, complete cds.  
 ACCESSION M27487  
 VERSION M27487.1 GI:703088  
 KEYWORDS cell surface glycoprotein; class II gene; integral membrane protein; lymphocyte antigen; major histocompatibility complex.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 818)  
 AUTHORS Young, J.A., Lindsay, J., Bodmer, J.G. and Trowsdale, J.

TITLE Epitope recognition by a DP alpha chain-specific monoclonal antibody (DP11.1) is influenced by the interaction between the DP alpha chain and its polymorphic DP beta chain partner

JOURNAL Hum. Immunol. 23 (1), 37-44 (1988)

MEDLINE 89053719

PUBMED 2461352

COMMENT On Mar 9, 1995 this sequence version replaced gi:341719.  
 Original source text: Homo sapiens cDNA to mRNA.

## FEATURES

source  
 1..818  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="6p21.3"  
 /cell\_line="HLICRF"  
 1..818  
 /cell\_type="B lymphoblast"  
 /gene="HLA-DPA1"  
 36..818  
 /gene="HLA-DPA1"  
 /codon\_start=1  
 /product="MHC class II DP3-alpha"  
 /protein\_id="AAA63220.1"  
 /db\_xref="GI:703089"  
 /db\_xref="GDB:G00-120-634"  
 /translation="MRERDMFHRVAVILRALSLAFILSLRGAGAKADHVSTYAAV  
 QTHRTGEFMEFDEDFYVLDKKETVHLEFGQAFSGAAGGLANLAINNNLN  
 TLIQRNHTQATNDPPETVTPKGFVBLGQNTLICHIDKFFPPVNLVTWICNGELVT  
 EGVAESLFLPTDYSFKFKHYLTFTVPSAEDFYDCRVEHWGSDQPLLKHWEAQEPTOMP  
 ETTETVLGALGLVLGLVGLIIVGTVLIIKSLRSGHDPRAQGTL"

36..128  
 /gene="HLA-DPA1"  
 /note="G00-120-634"  
 129..815  
 /gene="HLA-DPA1"  
 /product="MHC class II DP3-alpha"  
 /note="G00-120-634"

## sig\_peptide

## mat\_peptide

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 818;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||  
 DB 168 CAGACGCATAGACCAACAGG 187

## RESULT 28

LOCUS HSSBA1 1048 bp mRNA linear PRI 13-DEC-1996  
 DEFINITION Human mRNA for SB classII histocompatibility antigen alpha-chain.  
 ACCESSION X00457 K01506  
 VERSION X00457.1 GI:36405  
 KEYWORDS antigen; membrane protein; signal peptide.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1048)  
 AUTHORS Auffray, C., Lillie, J.W., Arnot, D., Grossberger, D., Kappes, D. and Strominger, J.L.

TITLE Isotypic and allelotypic variation of human class II histocompatibility antigen alpha-chain genes

JOURNAL Nature 308 (5957), 327-333 (1984)

MEDLINE 84168117

PUBMED 6584734

COMMENT On Nov 6, 2003 this sequence version replaced gi:188516.  
 Location/Qualifiers

## FEATURES

source  
 1..1048  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 <1..703  
 /codon\_start=2  
 /product="SB classII histocompatibility antigen alpha-chain"  
 /protein\_id="CAA25143.1"  
 /db\_xref="GI:758100"  
 /db\_xref="GOA:P20036"  
 /db\_xref="SWISS-PROT:P20036"  
 /translation="GAGAIKADHVSYIAFVQTHRTGEMFEFDEDFYVLDKKE  
 TYWHLEFGQAFSGAAGGLANLAINNNLTIQRNHTQATNDPPETVTPKGFVPS  
 LGQPTLICHIDKFFPPVNLVTWICNGELVTGVAESLFLPTDYSFKFKHYLTFTVPS  
 AEDFYDCRVEHWGLDQPLLKHWEAQEPTETVLGALGLVLGLVGLIIVGTVLII  
 KSLRSGHDPRAQGTL"

## CDS

## sig\_peptide

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## polyA\_site

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1048;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||  
 DB 53 CAGACGCATAGACCAACAGG 72

## RESULT 29

LOCUS I03088 1140 bp ss-DNA linear PAT 21-MAY-1993  
 DEFINITION Sequence 7 from Patent US 4582789.  
 ACCESSION I03088  
 VERSION I03088.1 GI:268244  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.

```

REFERENCE 1 (bases 1 to 1140)
AUTHORS Sheldon,E.L. III, Levenson,C.H., Mullis,K.B. and Rapoport,H.
TITLE Process for labeling nucleic acids using psoralen derivatives
JOURNAL Patent: US 4582789-A 7 15-APR-1986;
Cetus Corporation, Emeryville, CA
FEATURES
    source
        1..1140
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1140;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGACGCATAGACCAACAGG 20
Db 212 CAGACGCATAGACCAACAGG 231
E00484
LOCUS E00484 1201 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of pSBalpha-3.6.
ACCESSION E00484
VERSION E00484.1 GI:2168767
KEYWORDS JP 1985226888-A/2.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Edowado,R.S.Z., Kerii,B.M., Koorei,H.R. and Henrili,R.
TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL Patent: JP 1985226888-A 2 12-NOV-1985;
CETUS CORP
OS Human {Homo sapiens}
FN JP 1985226888-A/2
PD 12-NOV-1985
PF 20-MAR-1985 JP 1985054705
PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 P
EDOWADO RUISU SHLERUDON ZA SAADO, KSERII BANKUSU MIYURISU, PI
KOOREI HANAADO REBENSON, HENRII RAPOPORUTO
PC C07D519/00, C07H21/04, C12N15/00, C12Q1/68, G01N33/50, G01N33/532,
PC G01N33/58,
PC C07D519/00, C07D493/04, C07D495/04, C07D519/00, C07D493/04, PC
C07D493/10;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue type=Blood;
CC *source: cell type=B cell;
CC *source: library=cDNA library;
CC *source: clone=pSBalpha-318;
FH Key
FH Location/Qualifiers
FT CDS
FT 1..2201
FT /gene='pSBalpha-318'.
FEATURES
    source
        1..1201
            /organism="unidentified"
            /mol_type="genomic RNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGACGCATAGACCAACAGG 20
Db 212 CAGACGCATAGACCAACAGG 231

```

```

RESULT 31
LOCUS I03006 1201 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4617261.
ACCESSION I03006
VERSION I03006.1 GI:268462
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and
Watson,R.M.
TITLE Process for labeling nucleic acids and hybridization probes
JOURNAL Patent: US 4617261-A 6 14-OCT-1986;
Cetus Corporation, Emeryville, CA
FEATURES
    source
        1..1201
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGACGCATAGACCAACAGG 20
Db 212 CAGACGCATAGACCAACAGG 231
I03423
LOCUS I03423 1201 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4822731.
ACCESSION I03423
VERSION I03423.1 GI:270023
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Watson,R.M., Sheldon,E.L. III and Shead,R.M.
TITLE Process for labeling single-stranded nucleic acids and
hybridization probes
JOURNAL Patent: US 4822731-A 6 18-APR-1989;
Cetus Corporation, Emeryville, CA
FEATURES
    source
        1..1201
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGACGCATAGACCAACAGG 20
Db 212 CAGACGCATAGACCAACAGG 231
RESULT 33
LOCUS AX552229 1259 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 138 from Patent WO0162927.
ACCESSION AX552229
VERSION AX552229.1 GI:25896467
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 128.353 Seconds

(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819b-39

Perfect score: 20

Sequence: - cagacgcatacagcaacagg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

N Geneseq\_29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003as:\*

8: geneseqn2003bs:\*

9: geneseqn2003cs:\*

10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	6	Aal48203 Human HLA
2	20	100.0	21	6	Aal48204 Human HLA
3	20	100.0	255	6	Aal48219 Human HLA
4	20	100.0	267	4	Aak54418 Human hae
5	20	100.0	267	4	Aak54603 Human hae
6	20	100.0	272	4	Aak54317 Human hae
7	20	100.0	294	4	Aak54555 Human hae
8	20	100.0	294	4	Aak54805 Human hae
9	20	100.0	362	3	Aak43013 Human hae
10	20	100.0	410	8	ACH43248 Human hae
11	20	100.0	436	3	ACH43248 Human hae
12	20	100.0	456	2	AAV86216 EST c10re
13	20	100.0	476	3	AAC00117 Human sec
14	20	100.0	490	8	ACH49928 Human leu
15	20	100.0	576	3	AAC00116 Human sec
16	20	100.0	1202	2	AAC00116 Human sec
17	20	100.0	1259	4	AAC25060 pSbalpha-
18	20	100.0	1348	3	AAS31123 Human dia
19	20	100.0	14646	6	Aaf18332 Lung canc
20	19	95.0	4233	3	ABK64796 Human ben
21	19	95.0	4233	4	Aaz290198 Rat mdrlb
22	19	95.0	4254	6	Aaf27498 Rat mdrlb
23	19	95.0	4254	7	ABK63517 Rat seqe
					ABT41782 Toxicity

24	18.4	92.0	396	2	AAV86130	AAV86130 EST clone
25	17.4	87.0	96596	8	ADA02564	ADA02564 Human RAS
26	17.4	87.0	96596	9	ADB72302	ADB72302 Human RAS
27	17.4	87.0	240000	7	ACD13446	ACD13446 Human DNA
28	16.8	84.0	110000	4	AAI99682_03	Continuation (4 of
29	16.8	84.0	110000	4	AAI99683_03	Continuation (4 of
30	16.4	82.0	4189	3	AAZ49334	Aaz49334 Murine mu
31	16.4	82.0	4189	6	ABA94367	ABA94367 Mouse BCR
32	16.4	82.0	4298	9	ADD29583	Add29583 Mouse tum
33	16.4	82.0	4313	2	AAQ38950	Aaq38950 Mouse mul
34	16	80.0	279	2	AAZ32177	Aaz32177 Human coa
35	16	80.0	396	4	AAI82378	Aai82378 Human pol
36	16	80.0	148567	7	ABS55500	Abse55500 Gene enco
37	16	80.0	148567	8	ACA62841	ACA62841 Human kin
38	15.8	79.0	65	6	ABNS2705	ABNS2705 Mouse spl
39	15.8	79.0	267	2	AAV86047	AAV86047 EST clone
40	15.8	79.0	271	6	ABES1596	Abes1596 Human cdn
41	15.8	79.0	612	6	ABN63360	ABn63360 Human can
42	15.8	79.0	1183	6	ABL55410	ABl55410 Human Ll
43	15.8	79.0	26410	4	AAK70623	AAk70623 Human imm
44	15.8	79.0	40116	7	ABZ26080	ABz26080 Mouse DNA
45	15.8	79.0	96599	8	ADA02981	Ada02981 Mouse Map
46	15.8	79.0	96599	9	ADB72719	ADB72719 Mouse Map
47	15.8	79.0	96599	9	ADC85461	ADC85461 Mouse Map
48	15.8	79.0	110000	7	AD53224_2	Continuation (3 of
49	15.8	79.0	112132	6	ABK90888	ABK90888 Human ATP
50	15.8	79.0	148834	6	ABK83570	ABK83570 Human cDN
51	15.8	79.0	322101	9	AAAD58431	AAAD58431 Human PAC
52	15.4	77.0	565	8	ABV55407	ABv55407 Human pro
53	15.4	77.0	1068	8	ADA31143	Ada31143 DNA encod
54	15.4	77.0	1353	7	ACA23947	ACA23947 Prokaryot
55	15.4	77.0	27150	4	AAAD02701	AAAD02701 Human gly
56	15.4	77.0	54391	7	ACF42745_3	Continuation (4 of
57	15.4	77.0	172637	6	ABN83124	ABn83124 Human vol
58	15.4	77.0	237961	6	ABQ80552	ABq80552 Human Can
59	15.2	76.0	140	4	ABA66598	ABa66598 Human hae
60	15.2	76.0	140	4	AAK40756	AAk40756 Human bon
61	15.2	76.0	140	4	AAK15025	AAk15025 Human bra
62	15.2	76.0	140	4	ABSA0330	ABsA0330 Human liv
63	15.2	76.0	395	6	ABL66554	ABl66554 Lung canc
64	15.2	76.0	423	5	ABAI1047	ABaI1047 Human ner
65	15.2	76.0	439	6	ABN95367	ABn95367 Gene #186
66	15.2	76.0	446	5	ABL78941	ABl78941 Human ova
67	15.2	76.0	454	4	AAI08677	AAI08677 Human bre
68	15.2	76.0	477	4	ABAS4027	ABaS4027 Human foe
69	15.2	76.0	477	4	AAK27744	AAK27744 Human bon
70	15.2	76.0	477	4	AAK02300	AAk02300 Human bra
71	15.2	76.0	477	4	ABST7322	ABsT7322 Human liv
72	15.2	76.0	547	4	AAK78196	AAk78196 Human imm
73	15.2	76.0	547	4	AAK78198	AAk78198 Human imm
74	15.2	76.0	547	4	AAK78197	AAk78197 Human imm
75	15.2	76.0	666	4	AAI71573	AAI71573 Corynebact
76	15.2	76.0	672	4	AAI71574	AAI71574 Corynebact
77	15.2	76.0	912	4	AAH98478	AAh98478 Human EST
78	15.2	76.0	933	4	AAH98506	AAh98506 Human EST
79	15.2	76.0	1178	2	AAV69563	AAv69563 Soil deri
80	15.2	76.0	1356	3	AAZ53678	AAz53678 Neisseria
81	15.2	76.0	1449	5	AAAS65954	AAAS65954 DNA encod
82	15.2	76.0	1538	9	ADE09924	Ade09924 Novel DNA
83	15.2	76.0	1554	5	AAH67693	AAh67693 C glutami
84	15.2	76.0	1691	3	AAAS9769	AAAS9769 Human sec
85	15.2	76.0	1847	5	ABAI4426	ABaI4426 Human ner
86	15.2	76.0	2122	4	AAI26582	AAI26582 DNA encod
87	15.2	76.0	2122	7	ABZ73292	ABz73292 Secreted
88	15.2	76.0	2122	9	ADC20072	ADc20072 Human sec
89	15.2	76.0	2270	9	ADE54693	Ade54693 Human gen
90	15.2	76.0	2338	6	AAI49314	AAI49314 Rat Glut1
91	15.2	76.0	2338	6	AAI49313	AAI49313 Rat Glut1
92	15.2	76.0	2338	6	AAI49315	AAI49315 Rat Glut1
93	15.2	76.0	2571	9	ABDS3698	ABDS3698 Primary r
94	15.2	76.0	3175	4	AAI99170	AAI99170 Human exc
95	15.2	76.0	3175	4	AAAS32858	AAAS32858 Human gen
96	15.2	76.0	3175	5	AAI63520	AAI63520 Human kid

c 97 15.2 75.0 3465 5 ABV25341 Human pro  
 c 98 15.2 76.0 5060 4 AAK78298 Human imm  
 c 99 15.2 76.0 5064 4 AAK78297 Human imm  
 c 100 15.2 76.0 5064 2 AAV81446 Pig p105  
 c 101 15.2 76.0 11348 4 AAK80405 Human imm  
 c 102 15.2 76.0 11350 4 AAK80407 Human imm  
 c 103 15.2 76.0 12642 7 ABT31904 Human bre  
 c 104 15.2 76.0 12680 6 ABT07741 Breast ca  
 c 105 15.2 76.0 12680 9 ABT75518 Prostate  
 c 106 15.2 76.0 12716 6 ABN59334 Novel hum  
 c 107 15.2 76.0 17069 4 AAK83892 Human imm  
 c 108 15.2 76.0 17069 4 AAK83892 Human imm  
 c 109 15.2 76.0 17310 4 AAK83881 Human imm  
 c 110 15.2 76.0 17897 9 ADB53395 DRG1 geno  
 c 111 15.2 76.0 28634 4 AAS59539 Propionib  
 c 112 15.2 76.0 28634 7 ACF64468 Propionib  
 c 113 15.2 76.0 33552 4 AAS13655 Genomic D  
 c 114 15.2 76.0 33552 8 ADA02678 Human TOP  
 c 115 15.2 76.0 36592 9 ADB72416 Human TIO  
 c 116 15.2 76.0 188888 6 ABO75562 Human rel  
 c 117 15.2 76.0 34980 5 AAH68532 C glutami  
 c 118 15 75.0 110000 7 AAD53223\_0 Human chr  
 c 119 14.8 74.0 157 4 AAK68741 Human imm  
 c 120 14.8 74.0 191 7 AAL14652 Prokaryot  
 c 121 14.8 74.0 223 7 AAL15030 Prokaryot  
 c 122 14.8 74.0 239 7 AAL14760 Prokaryot  
 c 123 14.8 74.0 239 7 AAL14846 Prokaryot  
 c 124 14.8 74.0 248 7 AAL14671 Prokaryot  
 c 125 14.8 74.0 278 7 AAL14875 Prokaryot  
 c 126 14.8 74.0 303 7 AAL14728 Prokaryot  
 c 127 14.8 74.0 325 7 AAL14488 Prokaryot  
 c 128 14.8 74.0 354 7 AAL14724 Prokaryot  
 c 129 14.8 74.0 355 7 AAL14872 Prokaryot  
 c 130 14.8 74.0 456 7 AAL14500 Prokaryot  
 c 131 14.8 74.0 457 7 AAL15187 Prokaryot  
 c 132 14.8 74.0 570 7 AAL15253 Prokaryot  
 c 133 14.8 74.0 754 5 AAS86696 DNA encod  
 c 134 14.8 74.0 832 5 AAS86701 DNA encod  
 c 135 14.8 74.0 1062 4 AAH34142 Human col  
 c 136 14.8 74.0 1097 7 AAL35256 Prokaryot  
 c 137 14.8 74.0 1274 4 AAF63382 Arf promo  
 c 138 14.8 74.0 1284 7 AAL49506 Prokaryot  
 c 139 14.8 74.0 1329 3 AAL32059 Prokaryot  
 c 140 14.8 74.0 1332 3 AAL65960 E. coli p  
 c 141 14.8 74.0 1332 4 AAS56363 Salmonell  
 c 142 14.8 74.0 1332 4 AAS52605 E. coli D  
 c 143 14.8 74.0 1332 4 AAH84515 E. coli g  
 c 144 14.8 74.0 1332 5 AAH84296 Escherich  
 c 145 14.8 74.0 1332 7 AAL18726 Prokaryot  
 c 146 14.8 74.0 1332 7 AAL18726 Prokaryot  
 c 147 14.8 74.0 1332 7 AAS1678 Prokaryot  
 c 148 14.8 74.0 1369 3 AAL40089 Arabidops  
 c 149 14.8 74.0 1890 7 AAO5564 Streptoco  
 c 150 14.8 74.0 1890 7 ABX05091 S. pneumo  
 c 151 14.8 74.0 2532 9 ADD12658 Human CDN  
 c 152 14.8 74.0 2719 4 ADE55968 Human gen  
 c 153 14.8 74.0 3088 7 AAT05628 ADP ribos  
 c 154 14.8 74.0 3158 7 ABX78412 Rice stre  
 c 155 14.8 74.0 3685 9 ADB62147 Human cDN  
 c 156 14.8 74.0 3883 4 ABL15834 Drosophil  
 c 157 14.8 74.0 4077 7 AAD51580 Human str  
 c 158 14.8 74.0 4172 4 ABL12206 Drosophil  
 c 159 14.8 74.0 4256 9 ADD12681 Human CDN  
 c 160 14.8 74.0 4918 2 AAT94024 Rat canal  
 c 161 14.8 74.0 4918 6 ABK63671 Rat seque  
 c 162 14.8 74.0 4918 7 ABT41847 Toxicity  
 c 163 14.8 74.0 4918 9 ADB58127 Toxicity-  
 c 164 14.8 74.0 5165 4 AAL03700 Human rep  
 c 165 14.8 74.0 5165 4 ABAC07883 Human ova  
 c 166 14.8 74.0 10944 4 AAS46250 DNA encod  
 c 167 14.8 74.0 12304 4 AAL03699 Human rep  
 c 168 14.8 74.0 12304 4 AAL07882 Human ova  
 c 169 14.8 74.0 20986 2 AAV52187 Streptoco

Abv25341 Human pro  
 Aak78298 Human imm  
 Aak78297 Human imm  
 Aav81446 Pig p105  
 Aak80405 Human imm  
 Aak80407 Human imm  
 Abt31904 Human bre  
 Abt07741 Breast ca  
 Abt75518 Prostate  
 Abn59334 Novel hum  
 Aak83892 Human imm  
 Aak83892 Human imm  
 Aak83881 Human imm  
 Adb53395 DRG1 geno  
 Aas59539 Propionib  
 Acf64468 Propionib  
 Aas13655 Genomic D  
 Ada02678 Human TOP  
 Adb72416 Human TIO  
 Abo75562 Human rel  
 Aah68532 C glutami  
 Aad53223 Human chr  
 Aak68741 Human imm  
 Aal14652 Prokaryot  
 Aal15030 Prokaryot  
 Aal14760 Prokaryot  
 Aal14846 Prokaryot  
 Aal14671 Prokaryot  
 Aal14875 Prokaryot  
 Aal14728 Prokaryot  
 Aal14488 Prokaryot  
 Aal14724 Prokaryot  
 Aal14872 Prokaryot  
 Aal14500 Prokaryot  
 Aal15187 Prokaryot  
 Aal15253 Prokaryot  
 Aas86696 DNA encod  
 Aas86701 DNA encod  
 Aah34142 Human col  
 Aa35256 Prokaryot  
 Aaf63382 Arf promo  
 Aa49506 Prokaryot  
 Aa32059 Prokaryot  
 Aa65960 E. coli p  
 Aa56363 Salmonell  
 Aa52605 E. coli D  
 Aah84515 E. coli g  
 Aah8296 Escherich  
 Aal18726 Prokaryot  
 Aal18726 Prokaryot  
 Aa51678 Prokaryot  
 Aa40089 Arabidops  
 Aa05564 Streptoco  
 Abx05091 S. pneumo  
 Add12658 Human CDN  
 Ade55968 Human gen  
 Aat05628 ADP ribos  
 Abx78412 Rice stre  
 Adb62147 Human cDN  
 Abl15834 Drosophil  
 Aad51580 Human str  
 Abl12206 Drosophil  
 Add12681 Human CDN  
 Aat94024 Rat canal  
 Abk63671 Rat seque  
 Abt41847 Toxicity  
 Adb58127 Toxicity-  
 Aal03700 Human rep  
 Aba07883 Human ova  
 Aa46250 DNA encod  
 Aal03699 Human rep  
 Aal07882 Human ova  
 Aav52187 Streptoco

c 170 14.8 74.0 21188 9 ADC86606  
 c 171 14.8 74.0 28564 9 ADD48759  
 c 172 14.8 74.0 110000 3 AAF22305\_05  
 c 173 14.8 74.0 110000 7 ABS95454\_03  
 c 174 14.8 74.0 110000 7 ABX16390\_4  
 c 175 14.8 74.0 110000 8 ADB12062\_06  
 c 176 14.8 74.0 134141 6 ABN83487  
 c 177 14.8 74.0 349980 6 ABQ81842  
 c 178 14.4 72.0 157 3 AAC30054  
 c 179 14.4 72.0 168 3 AAC17921  
 c 180 14.4 72.0 343 3 AAC16444

Adc86606 Human GPC  
 Add48759 Human gen  
 Continuation (6 of  
 Continuation (4 of  
 Continuation (5 of  
 Continuation (7 of  
 Abn83487 Escherich  
 Abq81842 Bifidobac  
 Aac30054 Human sec  
 Aac17921 Human sec  
 Aac16444 Human sec

AAL48203  
 ID AAL48203 standard; DNA; 20 BP.  
 XX AAL48203;  
 AC AAL48203;  
 DT DT  
 01-OCT-2002 (first entry)  
 Human HLA DPA1 locus polymorphism primer sequence #3.  
 Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
 Flow cytometry; human; DPA1; PCR; primer; ss.  
 Homo sapiens.  
 W0200194639-A1.  
 13-DEC-2001.  
 07-JUN-2001; 2001WO-US018590.  
 08-JUN-2000; 2000US-0210759P.  
 (REGC ) UNIV CALIFORNIA.  
 White PS, Torney DC;  
 WPI; 2002-566450/60.  
 Identifying sequences useful as address/capture tags for flow cytometry  
 based minisequencing, by generating tag sequences and rejecting sequences  
 based on certain parameters e.g. sequences which form stable hairpins.  
 Disclosure; Page 14; 35pp; English.  
 The present invention relates to a method of identifying sequences useful  
 as addresses/capture tags, involving rejecting sequences having common sub-  
 sequences with a sub-sequence length greater than specified number of  
 bases, and sequences which can form stable hairpins and stable dimers  
 from a sample of oligonucleotides, and selecting those sequences in the  
 sample that would hybridize to their respective complements with a high  
 degree of specificity. The method is useful for identifying a set of  
 sequences useful as address/capture tags which can be used for  
 multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
 cytometry assay. The present sequence is a primer tag described in the  
 exemplification of the invention

SQ Sequence 20 BP; 8 A; 6 C; 5 G; 1 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACCATAGACCAACAGG 20  
 |||||  
 Db 1 CAGACCATAGACCAACAGG 20



XX		WO200194635-A1.
PN		XX
XX		13-DEC-2001.
PD		XX
PP		XX
PX		07-JUN-2001; 2001WO-US018590.
PF		XX
PG		XX
PH		08-JUN-2000; 2000US-0210759P.
PI		XX
PJ		(REGC ) UNIV CALIFORNIA.
PK		XX
PL		White PS, Torney DC;
PM		XX
PN		WPI; 2002-566450/60.
PO		XX
PP		XX
PQ		Identifying sequences useful as address/capture tags for flow cytometry based minisequencing, by generating tag sequences and rejecting sequences based on certain parameters e.g. sequences which form stable hairpins.
PR		XX
PS		Disclosure; Fig 4; 35pp; English.
PT		XX
PV		The present invention relates to a method of identifying sequences useful as address/capture tags, involving rejecting sequences having common sub-
PW		sequences with a sub-sequence length greater than specified number of bases, and sequences which can form stable hairpins and stable dimers from a sample of oligonucleotides, and selecting those sequences in the sample that would hybridize to their respective complements with a high degree of specificity. The method is useful for identifying a set of sequences useful as address/capture tags which can be used for multiplexed single nucleotide polymorphism (SNP) scoring in a flow cytometry assay. The present sequence is a fragment of the human DPA1 gene described in the exemplification of the invention
PY		XX
PZ		Sequence 255 BP; 67 A; 57 C; 66 G; 65 T; 0 U; 0 Other;
QA		XX
QB		Query Match 100.0%; Score 20; DB 6; Length 255;
QC		Best Local Similarity 100.0%; Pred.No.3;
QD		Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QE		QY 1 CAGACGCATAGACCACAGG 20 
QF		D5 40 CAGACGCATAGACCACAGG 59 
QG		RESULT 4
QH		AAK54418
QI		ID AAK54418 standard; cDNA; 267 BP.
QJ		XX AC AAK54418;
QK		XX AC
QL		XX
QM		DT 13-NOV-2001 (first entry)
QN		XX
QO		DE Human haematological malignancy-related antigen coding sequence #143.
QP		DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
QQ		KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
QR		KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
QS		XX Homo sapiens.
QT		OS
QU		XX WO200164886-A2.
QV		XX
QW		PD 07-SEP-2001.
QX		XX
QY		XX 01-MAR-2001; 2001WO-US007272.
RA		XX
RB		01-MAR-2000; 2000US-0186126P.
RC		PR 17-MAR-2000; 2000US-019047P.
RD		PR 27-APR-2000; 2000US-0200545P.
RE		PR 28-APR-2000; 2000US-0200303P.
RF		PR 28-APR-2000; 2000US-0200775P.
RG		PR 01-MAY-2000; 2000US-0200999P.
RH		PR 04-MAY-2000; 2000US-0202084P.

```

PR 22-MAY-2000; 2000US-0206201P.
PR 14-JUL-2000; 2000US-0218950P.
PR 03-AUG-2000; 2000US-0222903P.
PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
XX
XX
XX
XX Gaiger A, Algate PA, Mannion J;
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX Claim 31; Page 371; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human haematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of haematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Haematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 5
AAK54503
ID AAK54603 standard; cDNA; 267 BP.
XX
XX AAK54603;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #328.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO200164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US0007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX

```

```

PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX Claim 31; Page 418; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of haematological malignancies. The
XX present sequence is the coding sequence of a human haematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of haematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Haematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 6
AAK54317
ID AAK54317 standard; cDNA; 272 BP.
XX
XX AAK54317;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #42.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO200164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US0007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX

```

PA (CORI-) CORIXA CORP.  
 XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 31; Page 345; 1252pp; English.  
 XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of hematological malignancies. The  
 CC present sequence is the coding sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX Sequence 272 BP; 67 A; 60 C; 75 G; 70 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 20; DB 4; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGACGGCATAGACCAACAGG 20  
 Db |||||  
 166 CAGACGGCATAGACCAACAGG 185  
 RESULT 7  
 AAK54555/c  
 ID AAK54555 standard; cDNA; 294 BP.  
 AC AAK54555;  
 XX 13-NOV-2001 (first entry)  
 DT Human haematological malignancy-related antigen coding sequence #280.  
 DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 XX haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
 XX Homo sapiens.  
 OS WO200164886-A2.  
 PN 07-SEP-2001.  
 PD 01-MAR-2001; 2001WO-US007272.  
 XX 01-MAR-2000; 2000US-0186126P.  
 PF 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200779P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 01-MAY-2000; 2000US-020084P.  
 PR 14-JUL-2000; 2000US-0218950P.  
 PR 03-AUG-2000; 2000US-0206201P.  
 PR 22-MAY-2000; 2000US-0218950P.  
 PR 14-JUL-2000; 2000US-0206201P.  
 PR 03-AUG-2000; 2000US-0218950P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX (CORI-) CORIXA CORP.  
 PA Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX

XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 31; Page 406; 1252pp; English.  
 XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the coding sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 20; DB 4; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGACGGCATAGACCAACAGG 20  
 Db |||||  
 107 CAGACGGCATAGACCAACAGG 88  
 RESULT 8  
 AAK54805/c  
 ID AAK54805 standard; cDNA; 294 BP.  
 AC AAK54805;  
 XX 13-NOV-2001 (first entry)  
 DT Human haematological malignancy-related antigen coding sequence #530.  
 DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
 XX Homo sapiens.  
 OS WO200164886-A2.  
 PN 07-SEP-2001.  
 PD 01-MAR-2001; 2001WO-US007272.  
 XX 01-MAR-2000; 2000US-0186126P.  
 PF 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 28-APR-2000; 2000US-020084P.  
 PR 14-JUL-2000; 2000US-0218950P.  
 PR 03-AUG-2000; 2000US-0206201P.  
 PR 22-MAY-2000; 2000US-0218950P.  
 PR 14-JUL-2000; 2000US-0206201P.  
 PR 03-AUG-2000; 2000US-0218950P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX (CORI-) CORIXA CORP.  
 PA Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX

PT Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.

PS Claim 31; Page 469; 1252pp; English.

XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of hematological malignancies. The  
 CC present sequence is the coding sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of hematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma

SQ Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 294;

Best Local Similarity 100.0%; Pred. No. 3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

|||||

107 CAGACGCATAGACCAACAGG 88

RESULT 9

AAA43013

ID AAA43013 standard; cDNA; 362 BP.

XX AAA43013;

21-AUG-2000 (first entry)

Human secreted expressed sequence tag SEQ ID NO:1753.

XX Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiaesthetic; vulnerary; antiparkinsonian;  
 KW antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
 KW infection; depression; psoriasis; ss.

OS Homo sapiens.

XX WO200021990-A1.

PN 20-APR-2000.

PF 15-OCT-1999; 99WO-US024205.

XX 15-OCT-1998; 98US-010443EP.

PR (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M;

XX WPI; 2000-317937/27.

PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (SESTs), useful for treating various disorders

PT such as autoimmune, infectious, and central nervous system disorders.

XX Claim 1; Page 526; 618pp; English.

XX AAA41261 to AAA43419 represent specifically claimed secreted expressed  
 CC sequence tags (SESTs), isolated from human, mouse, xenopus and rat tissue  
 CC sources. The SESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
 CC antiviral; antidiabetic; antiaesthetic; vulnerary; antitumor;  
 CC osteoprotective; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;  
 CC cerebroprotective; anticonvulsant; and antidepressant. The SESTs can be  
 CC used for gene therapy and in vaccines. The SESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention

SQ Sequence 362 BP; 86 A; 86 C; 99 G; 91 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 362;

Best Local Similarity 100.0%; Pred. No. 3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

|||||

165 CAGACGCATAGACCAACAGG 184

RESULT 10

ACH49248

ID ACH49248 standard; cDNA; 410 BP.

XX ACH49248;

13-OCT-2003 (first entry)

XX Human leukocyte cDNA #842.

XX Human; ss; sequencing by hybridisation; SH4; expressed sequence tag; SEST;  
 KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful  
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene

PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 36460; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
XX determined by the technique of SBH (sequencing by hybridisation). Also  
XX included is a purified polypeptide comprising a sequence corresponding to  
XX a reading frame of the novel polynucleotide. The nucleic acid sequences  
XX are useful in diagnostics as expressed sequence tags (EST) for  
XX identifying expressed genes or for physical mapping of the human genome,  
XX in forensics, in assessing biodiversity, or in identifying mutations  
XX responsible for genetic disorders and other traits. The nucleotide  
XX sequences are also useful as hybridisation probes, as oligomers for PCR,  
XX for chromosome and gene mapping, in the recombinant production of  
XX protein, or in generating antisense DNA or RNA. The purified polypeptide  
XX is useful for generating antibodies specific for it. The present sequence  
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 20; DB 8; Length 410;

XX Best Local Similarity 100.0%; Pred. No. 3.1; Mismatches 0; Indels 0; Gaps 0;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

DB 184 CAGACGCATAGACCAACAGG 203

RESULT 11

AAA43818

ID AAA43818 standard; cDNA; 436 BP.

XX AAA43818;

DT 21-AUG-2000 (first entry)

XX Mouse secreted expressed sequence tag SEQ ID NO:393.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;  
XX expressed sequence tag; EST; probe; chemotactic; proliferative;  
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
XX antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;  
XX antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
XX autoimmune disorder; multiple sclerosis; allergic condition; ulcer;  
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
XX central nervous system disorder; Alzheimer's disease; stroke;  
XX Parkinson's disease; Huntington's disease; coagulation disorder;  
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
XX infection; depression; psoriasis; ss.

OS Mus musculus.

XX WO2000021991-A1.

XX 20-APR-2000.

XX 15-OCT-1999; 99WO-US024206.

XX 15-OCT-1998; 98US-0104436P.

XX (GEMY ) GENETICS INST INC.

XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Bowman MR;

XX DR

XX WPI; 2000-317938/27.

XX Isolated polynucleotides, and encoded proteins, comprising secreted  
XX expressed sequence tags (sESTs), useful for treating various disorders  
XX such as autoimmune, infectious, and central nervous system disorders.

XX Claim 1; Page 306; 803pp; English.

XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
XX sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue  
XX sources. The sESTs can have a range of activities depending on the  
XX tissues they were isolated from. The activities include: chemotactic;  
XX proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
XX haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
XX antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antitumor;  
XX osteoprotective; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;  
XX cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be  
XX used for gene therapy and in vaccines. The sESTs are useful as probes for  
XX the identification and isolation of full-length cDNAs and genomic DNA  
XX molecules which correspond to the sESTs. Proteins encoded by the sESTs  
XX are useful in assays for determining biological activity and raising  
XX antibodies. They may be useful for treatment of autoimmune disorders  
XX (multiple sclerosis, insulin dependent diabetes), allergic conditions  
XX (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
XX osteoporosis, osteoarthritis, central nervous system disorders  
XX (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
XX disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
XX disease), tumours, bacterial, fungal or viral infections, depression and  
XX psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
XX in the exemplification of the present invention

XX Sequence 436 BP; 108 A; 113 C; 106 G; 109 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 20; DB 3; Length 436;

XX Best Local Similarity 100.0%; Pred. No. 3.1;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

DB 206 CAGACGCATAGACCAACAGG 225

RESULT 12

AAV66216

ID AAV66216 standard; cDNA; 466 BP.

XX AAV66216;

DT 27-APR-1999 (first entry)

XX EST clone 067.

XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
XX tissue growth; activin; inhibitor; tumour invasion suppressor; EST; human;  
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

OS Homo sapiens.

XX WO9845435-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98WO-US006954.

XX 10-APR-1997; 97US-00835913.

XX (GEMY ) GENETICS INST INC.

XX Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

PI Spaulding V, Agostino MJ;

XX

DR WPI; 1999-070076/06.  
 XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 XX  
 PS Claim 1; Page 160-161; 633pp; English.  
 XX  
 CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 466 BP; 110 A; 122 C; 122 G; 112 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 2; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 3.1;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 152 CAGACGCATAGACCAACAGG 171  
 RESULT 13  
 AAC00117  
 ID AAC00117 standard; cDNA; 476 BP.  
 AC AAC00117;  
 XX  
 XX 06-OCT-2000 (first entry)  
 XX Human secreted protein 5' EST, SEQ ID NO: 115.  
 DE  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 XX RP1033401-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX 21-FEB-2000; 2000EP-00200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487P.  
 PR  
 XX (GSET ) GENSET.  
 PA  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR P-PSDB; AAC00111.  
 DR  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX  
 PS Claim 1; SEQ ID NO 115; 71pp + Sequence Listing; English.  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA- RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors  
 XX  
 SQ Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 3; Length 476;  
 Best Local Similarity 100.0%; Pred. No. 3.1;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 227 CAGACGCATAGACCAACAGG 246  
 RESULT 14  
 ACH49928  
 ID ACH49928 standard; cDNA; 490 BP.  
 XX  
 AC ACH49928;  
 XX  
 XX 13-OCT-2003 (first entry)  
 XX Human leukocyte cDNA #1522.  
 DE  
 XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 XX Homo sapiens.  
 XX US2003073623-A1.  
 PN  
 XX 17-APR-2003.  
 PD  
 XX 30-JUL-2001; 2001US-00918995.  
 PF  
 XX 30-JUL-2001; 2001US-00918995.  
 PR  
 XX (DRMA/) DRMANAC R T.  
 PA (LABA/) LABAT I.  
 PA (STAC/) STACHE-CRAIN B.  
 PA (DICK/) DICKSON M C.  
 PA (JONE/) JONES L W.  
 XX  
 XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
 PI WPI; 2003-615964/58.  
 DR  
 XX New polynucleotide sequences obtained from various cDNA libraries, useful  
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
 PT mapping, in the recombinant production of protein, or in generating  
 PT antisense DNA or RNA.  
 XX  
 PS Claim 1; SEQ ID NO 37140; 44pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide comprising any one of  
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 CC determined by the technique of SBH (sequencing by hybridisation). Also  
 CC included is a purified polypeptide comprising a sequence corresponding to  
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
 CC are useful in diagnostics as expressed sequence tags (EST) for  
 CC identifying expressed genes or for physical mapping of the human genome,  
 CC in forensics, in assessing biodiversity, or in identifying mutations  
 CC responsible for genetic disorders and other traits. The nucleotide  
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
 CC for chromosome and gene mapping, in the recombinant production of

CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
 CC is useful for generating antibodies specific for it. The present sequence  
 CC is one of the 3843 isolated cDNA/EST sequences. Note: the sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 490 BP; 112 A; 130 C; 125 G; 121 T; 0 U; 2 Other;

Query Match 100.0%; Score 20; DB 8; Length 490;

Best Local Similarity 100.0%; Pred. No. 3.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 257 CAGACGCATAGACCAACAGG 276

RESULT 15

AAC00116

ID AAC00116 standard; cDNA; 576 BP.

XX AAC00116;

AC AAC00116;

XX 06-OCT-2000 (first entry)

DT Human secreted protein 5' EST, SEQ ID NO: 114.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

KW Homo sapiens.

OS Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00110.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 114; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from

XX mRNAs encoding secreted proteins. An ORF has been identified within the

XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

XX derived from 30 different tissues. EST sequences usually correspond

XX mainly to the 3' untranslated region (UTR) of the mRNA because they are

XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

XX well suited for isolating cDNA sequences derived from the 5' ends of

XX mRNAs and even in those cases where longer cDNA sequences have been

XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

XX cDNAs with intact 5' ends and can therefore be used to obtain full length

XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

XX gene therapy and chromosome mapping procedures. They are used to obtain

XX upstream regulatory sequences and to design expression and secretion

XX vectors

XX Sequence 576 BP; 131 A; 158 C; 141 G; 145 T; 0 U; 1 Other;

Query Match 100.0%; Score 20; DB 3; Length 576;

Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 327 CAGACGCATAGACCAACAGG 346

RESULT 16

AAQ25060

ID AAQ25060 standard; DNA; 1202 BP.

XX AAQ25060;

AC AAQ25060;

XX 15-NOV-1992 (first entry)

DT pSAlpha-318 clone.

DE RFLP; probe p29G8; beta-lymphoblastoid cell line; LG2; T5-1; 6.3.6;

XX transplant; transfection; paternity; ss.

KW Synthetic.

OS US5110920-A.

XX 05-MAY-1992.

XX 05-DEC-1984; 84US-00678255.

XX 22-JAN-1982; 82US-00341902.

XX 07-JAN-1983; 83US-00456373.

XX 30-AUG-1988; 88US-00238619.

XX (CETU ) CETUS CORP.

XX Erlich HA;

XX WPI; 1992-175244/21.

XX New DNA probes specific to single class II HLA locus - useful in HLA

XX typing e.g. to evaluate paternity and transplant or transfection

XX compatibility and to diagnose disease susceptibility.

XX Disclosure; Page 11; 21pp; English.

XX The sequence given is a pSAlpha-318 clone which was derived from a beta-

XX lymphoblastoid cell line LG2 cDNA library using a probe designated p29G8.

XX This probe bound to sequences distinct from those which lead to the

XX elucidation of HLA-Dp34 (see also AAQ25059). A genomic blot pattern with

XX DNA from the cell lines T5-1 and its HLA hemizygous derivative 6.3.6

XX indicates that the p29G8 locus maps within the HLA region. p29G8 has been

XX found to be a HLA-SAlpha clone and could be used to isolate the given

XX sequence. p29G8 could be useful in HLA typing based on RFLPs. It can be

XX utilized in paternity disputes or for determining transplant or transfection

XX compatibility. It can also be used to make disease correlations to

XX diagnose diseases or predict susceptibility to diseases

XX Sequence 1202 BP; 289 A; 336 C; 277 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 1202;

Best Local Similarity 100.0%; Pred. No. 3.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 212 CAGACGCATAGACCAACAGG 231

RESULT 17

AAS31123

ID AAS31123 standard; cDNA; 1259 BP.

XX AAS31123;

XX AAS31123;

DT 04-DEC-2001 (first entry)  
 XX Human diagnostic and therapeutic polynucleotide (DITHP) #138.  
 DE Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
 XX cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
 KW respiratory disorder; ss.  
 XX Homo sapiens.  
 OS W0200162927-A2.  
 XX 30-AUG-2001.  
 PD 21-FEB-2001; 2031W0-US006059.  
 XX 24-FEB-2000; 2000US-0184693P.  
 PR 24-FEB-2000; 2000US-0184697P.  
 PR 24-FEB-2000; 2000US-0184698P.  
 PR 24-FEB-2000; 2000US-0184768P.  
 PR 24-FEB-2000; 2000US-0184769P.  
 PR 24-FEB-2000; 2000US-0184770P.  
 PR 24-FEB-2000; 2000US-0184771P.  
 PR 24-FEB-2000; 2000US-0184772P.  
 PR 24-FEB-2000; 2000US-0184773P.  
 PR 24-FEB-2000; 2000US-0184774P.  
 PR 24-FEB-2000; 2000US-0184776P.  
 PR 24-FEB-2000; 2000US-0184777P.  
 PR 24-FEB-2000; 2000US-0184797P.  
 PR 24-FEB-2000; 2000US-0184813P.  
 PR 24-FEB-2000; 2000US-0184837P.  
 PR 24-FEB-2000; 2000US-0184841P.  
 PR 24-FEB-2000; 2000US-0185213P.  
 PR 24-FEB-2000; 2000US-0185216P.  
 PR 12-MAY-2000; 2000US-0203785P.  
 PR 15-MAY-2000; 2000US-0204226P.  
 PR 16-MAY-2000; 2000US-0204525P.  
 PR 16-MAY-2000; 2000US-0204821P.  
 PR 16-MAY-2000; 2000US-0204908P.  
 PR 16-MAY-2000; 2000US-0205232P.  
 PR 17-MAY-2000; 2000US-0204815P.  
 PR 17-MAY-2000; 2000US-0204863P.  
 PR 17-MAY-2000; 2000US-0205221P.  
 PR 17-MAY-2000; 2000US-0205285P.  
 PR 17-MAY-2000; 2000US-0205286P.  
 PR 17-MAY-2000; 2000US-0205287P.  
 PR 17-MAY-2000; 2000US-0205323P.  
 PR 17-MAY-2000; 2000US-0205324P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX Panzer SR, Spiro PA, Hanville SC, Shah P, Chalup MS, Chang SC;  
 PI Chen A, D'sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE, Dufour GE;  
 PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TF;  
 PI Roseberry AM, Rosen BH, Russo ED, Stockdreher TK, Daffo A;  
 PI Wright RJ, Yap PH, Yu JY, Bradley Dr, Bratcher SR, Chen W;  
 PI Cohen HZ, Hodgson DM, Lincoln SE, Jackson S;  
 XX WPI; 2001-502867/55.  
 DR P-PSDB; AAU19552.  
 XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 XX Claim 1; Page 361; 522pp; English.  
 PS The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DITHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide

CC expression by rectifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DITHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DITHPs and in assays to identify modulators of DITHP  
 CC expression and activity. The anti-DITHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DITHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant assay  
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic  
 CC (DITHP) polynucleotides of the invention  
 XX  
 SQ Sequence 1259 BP; 293 A; 350 C; 300 G; 316 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 4; Length 1259;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGACGCATAGACCAACAGG 20  
 Db 347 CAGACGCATAGACCAACAGG 366  
 RESULT 18  
 AAF18332  
 ID AAF18332 standard; DNA; 1348 BP.  
 XX  
 AC AAF18332;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polynucleotide sequence SEQ ID 351.  
 XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.  
 XX Homo sapiens.  
 OS  
 XX WO200055180-A2.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 08-MAR-2000; 2000WO-US005918.  
 PR  
 XX 12-MAR-1999; 99US-0124270P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Ruben SM;  
 XX  
 XX WPI; 2000-587514/55.  
 DR P-PSDB; AAB58456.  
 XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer.  
 XX Claim 1; Page 808-809; 1425pp; English.  
 PS Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer



CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytoskeletal; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerrary; gastrointestinal  
 CC general; nephrotropic; antineoplastic; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the protein  
 CC or polynucleotide sequences. The lung cancer associated polynucleotide  
 CC sequences may be used for detection of lung cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The proteins may be used to treat disorders such as  
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,  
 CC cardiovascular, renal, and proliferative disorders. The proteins may also  
 CC be used in the treatment of wounds and infectious diseases.  
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are  
 CC used in the course of the invention for the identification and  
 CC characterisation of the polynucleotide and protein sequences

XX  
 SQ Sequence 1348 BP; 335 A; 373 C; 307 G; 324 T; 0 U; 9 Other;  
 Query Match 100.0%; Score 20; DB 3; Length 1348;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 DB 285 CAGACGCATAGACCAACAGG 304

RESULT 19  
 ABK64796  
 ID ABK64796 standard; DNA; 14646 BP.  
 XX  
 AC ABK64796;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Human benign prostatic hyperplasia gene #691.  
 XX  
 KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 EN WO200212440-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 07-AUG-2001; 2001WO-US024708.  
 XX  
 PR 07-AUG-2000; 2000US-0223323P.  
 XX  
 PR 05-JUN-2001; 2001US-00873319.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 PA (NIBS) JAPAN TOBACCO INC.  
 XX  
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
 XX  
 WPI; 2002-257476/30.

XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 XX detecting expression levels of one or more genes in prostate cells from  
 XX patient that are differentially regulated compared to normal prostate  
 XX cells.

XX Disclosure; Page 373-377; 444pp; English.

XX The invention relates to a method of diagnosing (I) the onset or  
 XX progression of benign prostatic hyperplasia (BPH), or screening (II) for  
 XX or identifying an agent that modulates the onset or progression of BPH.  
 XX The method is based on changes in gene expression in BPH tissue isolated  
 XX from patients exhibiting different clinical states of prostate  
 XX hyperplasia as compared to normal prostate tissue. (I) comprises  
 XX detecting the expression levels of one or more genes in prostate cells  
 XX from the subject that are differentially regulated compared to normal  
 XX prostate cells. (II) comprises preparing a first gene expression profile

CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles. (I)  
 CC is useful for diagnosing the onset or progression of BPH. (II) is useful  
 CC for identifying an agent that modulates the onset or progression of BPH.  
 CC The methods are useful to present information identifying the expression  
 CC level in a tissue or cells, by comparing the expression level of genes  
 CC given in the specification in the tissue or cells to the level of  
 CC expression of gene in the database, and displaying the expression levels  
 CC of at least one gene in the tissue or cell sample compared to the  
 CC expression level in BPH. Agents using (II) are useful for treating BPH or  
 CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic  
 CC hyperplasia gene sequences of the invention.

XX Sequence 14646 BP; 4552 A; 2957 C; 2741 G; 4395 T; 0 U; 1 Other;

Query Match 100.0%; Score 20; DB 6; Length 14646;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 DB 6366 CAGACGCATAGACCAACAGG 6385

RESULT 20  
 AAZ90198/C  
 ID AAZ90198 standard; DNA; 4233 BP.  
 XX  
 AC AAZ90198;  
 XX  
 DT 19-MAY-2000 (first entry)  
 XX  
 DE Rat mdrlb2 (multispecific drug transporter) nucleotide sequence.  
 XX  
 KW Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;  
 KW cardiovascular disease; central nervous system disorder;  
 KW autoimmune disease; kidney disease; drug formulation; se.

XX Rattus sp.  
 OS  
 XX US6025160-A.  
 XX  
 PD 15-FEB-2000.  
 XX  
 PF 22-JUL-1998; 98US-00120513.  
 XX  
 PR 22-JUL-1998; 98US-00120513.  
 XX  
 PA (SMIK) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Brun KA, Ellens H, Yue L, Chenery RU, Feild JA;  
 XX  
 WPI; 2000-181810/16.  
 XX  
 P-PSDB; AAY78879.

XX Isolated polynucleotide encoding a rat mdrlb2 drug transporter  
 XX polypeptide, useful for treatment of e.g. cancer, autoimmune disease,  
 XX central nervous system disorders.

XX Claim 7; Col 13-17; 17pp; English.

XX This sequence represents a polynucleotide encoding a 1275 amino acid rat  
 XX mdrlb2 multispecific drug transporter polypeptide. The mdrlb2 nucleotide  
 XX sequence has cytostatic, antiinflammatory, cardiact, neuroprotective,  
 XX immunosuppressive and nephrotropic activity. Understanding the  
 XX functioning of the mdrlb2 polynucleotide and protein in transgenic animal  
 XX models is useful for treating and preventing diseases such as cancer,  
 XX inflammation, cardiovascular disease, central nervous system disorders,  
 XX autoimmune disease, and kidney disease. The use of the protein in cell  
 XX based, membrane based, or binding assays may enhance drug formulation,  
 XX selection of formulation excipients and compound design

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 U; 0 Other;  
Query Match 95.0%; Score 19; DB 3; Length 4233;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAGG 20  
|||||  
DB 946 AGACGCATAGACCAACAGG 928

RESULT 21  
AAF27498/c  
ID AAF27498 standard; cDNA; 4233 BP.

XX AC AAF27498;

XX DT 25-APR-2001 (first entry)

XX DE Rat mdrlb2 multidrug resistance cDNA, SEQ ID NO:1.

XX KW Rat; multidrug resistance; mdrlb2; multi-drug transporter family;  
XX KW drug screening; pharmacokinetic analysis; oral absorption;  
XX KW formulation design; bioavailability; transgenic animal; knockout animal;  
XX KW inflammation; cardiovascular disease; central nervous system disorder;  
XX KW cancer; autoimmune disorder; kidney disease; ss.

XX OS Rattus sp.

XX PN US6169166-B1.

XX PD 02-JAN-2001.

XX PF 29-NOV-1999; 99US-00450105.

XX PR 22-JUL-1998; 98US-00120513.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX PI Chenery RJ, Ellens H, Brun KA, Yue L, Feild JA;

XX WPI; 2001-158280/16.

XX DR P-PSDB; AAB60409.

XX PT New rat multidrug resistance protein (mdrlb2), useful for screening or  
XX PT identifying compounds that are (ant)agonists/inhibitors of the ratmdrlb2,  
XX PT as well as compounds with optimal development characteristics.

XX PS Disclosure; Col 13-14; 13pp; English.

XX CC The invention relates to a novel rat multidrug resistance (mdr) protein,  
XX CC mdrlb2 (AAB60409). Multi-specific drug transporter family proteins are  
XX CC present in cell which have a barrier function, such as intestinal  
XX CC epithelial cells, brain microvessel endothelial cells, kidney epithelial  
XX CC cells, and liver hepatocytes, and are also expressed by certain cancer  
XX CC cells. The rat mdrlb2 protein is useful for screening or identifying  
XX CC compounds that are agonists or antagonists of mdrlb2 activity. It may  
XX CC also be used to establish assays to predict oral absorption and  
XX CC pharmacokinetics of drugs in humans, and thus enhance the design of  
XX CC formulations through the identification of compounds with optimal  
XX CC development characteristics (i.e., high oral bioavailability, and once a  
XX CC day) dosing, reduced drug interactions, reduced variability, and reduced  
XX CC food effects), specifically to avoid interactions with human mdr-1.  
XX CC Transgenic and knockout animals created using DNA encoding the rat mdrlb2  
XX CC may be used to gain an insight into treating and preventing human  
XX CC diseases such as cancer, inflammation, cardiovascular disease, central  
XX CC nervous system disorders, autoimmune disorders and kidney disease. The  
XX CC present sequence represents cDNA encoding rat mdrlb2

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 U; 0 Other;

Query Match 95.0%; Score 19; DB 4; Length 4233;

Best Local Similarity 100.0%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAGG 20  
|||||  
DB 946 AGACGCATAGACCAACAGG 928

RESULT 22

ABK63517/c

ID ABK63517 standard; cDNA; 4254 BP.

XX AC ABK63517;

XX DT 18-JUN-2002 (first entry)

XX DE Rat sequence differentially expressed in response to a hepatotoxin #1424.

XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
XX KW differential expression; centrilobular necrosis; steatosis.  
XX OS Rattus norvegicus.

XX PN WO200210453-A2.

XX PD 07-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US023872.

XX PR 31-JUL-2000; 2000US-0222040P.

XX PR 02-NOV-2000; 2000US-0244880P.

XX PR 11-MAY-2001; 2001US-0290029P.

XX PR 15-MAY-2001; 2001US-0290645P.

XX PR 22-MAY-2001; 2001US-0292336P.

XX PR 06-JUN-2001; 2001US-0295798P.

XX PR 13-JUN-2001; 2001US-0297457P.

XX PR 19-JUN-2001; 2001US-0298884P.

XX PR 09-JUL-2001; 2001US-0303459P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

XX WPI; 2002-241625/29.

XX PT Predicting toxic effects of compounds or the progression of these toxic  
XX PT effects by determining the changes in gene expression in tissues or cells  
XX PT exposed to the toxin and comparing these to gene expression in unexposed  
XX PT tissues or cells.

XX PS Claim 1; SEQ ID NO 1424; 239pp; English.

XX CC The invention relates to methods for predicting toxic effects of  
XX CC compounds or the progression of these toxic effects. By determining the  
XX CC global changes in gene expression in tissues or cells exposed to the  
XX CC toxin and comparing these to gene expression in unexposed tissues or  
XX CC cells. Also included are methods of predicting at least one toxic effect  
XX CC of a compound or progression of a toxic effect, preferably the  
XX CC hepatotoxicity of a compound, comprising detecting the level of  
XX CC expression in a tissue or cell sample exposed to the compound of two or  
XX CC more genes listed in the specification, where differential expression of  
XX CC the genes is indicative of at least one toxic effect or progression. The  
XX CC method can also be used to identify an agent which modulates the toxic  
XX CC response and predict cellular pathways that a compound modulates in a  
XX CC cell. The methods utilize a set of at least two probes (on a solid  
XX CC support in kit form), where each of the probes comprises a sequence that  
XX CC specifically hybridizes to a gene listed in the specification, a computer  
XX CC system comprising a database containing information identifying the  
XX CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
XX CC set of genes comprising at least two genes listed in the specification,  
XX CC and a user interface to view the information used to present information  
XX CC identifying the expression level in a tissue or cell of at least one gene  
XX CC listed in the specification. The method is useful for elucidating global

CC changes in gene expression and for identifying toxicity markers in  
CC tissues or cell exposed to a known toxin. The genes may be used as  
CC toxicity markers in drug screening and toxicity assays. The genes and  
CC gene expression information may be used as diagnostic markers for the  
CC prediction or identification of the physiological state of tissue or cell  
CC sample that has been exposed to a compound or agent. Hepatotoxicity is  
CC characterized by centrilobular necrosis and steatosis. The present  
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
CC which is differentially expressed in response to a hepatotoxic agent  
XX  
SQ Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 U; 0 Other;  
Query Match 95.0%; Score 19; DB 6; Length 4254;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AGACGCATAGACCAACAGG 20  
Db 1026 AGACGCATAGACCAACAGG 1008  
RESULT 23  
ABT41782/c  
ID ABT41782 standard; DNA; 4254 BP.  
XX  
AC ABT41782;  
XX  
DT 26-JUN-2003 (first entry)  
XX  
DE Toxicity modelling related rat gene SRQ ID No 1484.  
XX  
KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;  
KW database; drug screening; toxicity assay; rat; ds.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200295000-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 22-MAY-2002; 2002WO-US016173.  
XX  
PR 22-MAY-2001; 2001US-0290335P.  
PR 13-JUN-2001; 2001US-0297523P.  
PR 19-JUN-2001; 2001US-0298925P.  
PR 10-JUL-2001; 2001US-0303807P.  
PR 10-JUL-2001; 2001US-0303808P.  
PR 10-JUL-2001; 2001US-0303810P.  
PR 28-AUG-2001; 2001US-0315047P.  
PR 27-SEP-2001; 2001US-0324928P.  
PR 22-OCT-2001; 2001US-0330462P.  
PR 01-NOV-2001; 2001US-0330867P.  
PR 21-NOV-2001; 2001US-0331805P.  
PR 06-DEC-2001; 2001US-0336144P.  
PR 19-DEC-2001; 2001US-0340873P.  
PR 21-FEB-2002; 2002US-0357842P.  
PR 21-FEB-2002; 2002US-0357843P.  
PR 21-FEB-2002; 2002US-0357844P.  
PR 15-MAR-2002; 2002US-0364134P.  
PR 08-APR-2002; 2002US-0370144P.  
PR 08-APR-2002; 2002US-0370206P.  
PR 08-APR-2002; 2002US-0370247P.  
PR 17-APR-2002; 2002US-0372794P.  
PR 21-APR-2002; 2002US-0371679P.  
XX  
(GENE-) GENE LOGIC INC.  
XX  
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;  
PI WPI; 2003-148464/14.  
XX  
XX  
XX Predicting at least one toxic effect of a compound, useful for toxicity  
PT modeling, comprises preparing a gene expression profile of a tissue or

PT cell sample exposed to the compound, and comparing the gene expression  
PT profile to a database.  
XX Example 4; Page; 446pp; English.  
XX  
CC The invention relates to a novel method of predicting at least one toxic  
CC effect of a compound. The method comprises a gene expression profile of a  
CC tissue or cell sample exposed to the compound, and comparing the gene  
CC expression profile to a database comprising at least part of the data or  
CC information given in the specification. The methods are useful for  
CC predicting at least one toxic effect of a compound, predicting the  
CC progression of a toxic effect of a compound, predicting the renal  
CC toxicity of a compound, or identifying toxicity markers in tissues or  
CC cells exposed to known renal toxin. The genes are useful as toxicity  
CC markers in drug screening and toxicity assays, in monitoring disease or  
CC physiological states, or disease progression. This polynucleotide  
CC represents a rat DNA sequence relating to the toxic effect database  
CC described in the specification. NOTE: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the World Intellectual Property  
CC Organization  
XX  
SQ Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 U; 0 Other;  
Query Match 95.0%; Score 19; DB 7; Length 4254;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AGACGCATAGACCAACAGG 20  
Db 1026 AGACGCATAGACCAACAGG 1008  
RESULT 24  
AAV86130  
ID AAV86130 standard; cDNA; 396 BP.  
XX  
AC AAV86130;  
XX  
DT 27-APR-1999 (first entry)  
XX  
DE EST clone H45.  
XX  
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO9845435-A2.  
XX  
PD 15-OCT-1998.  
XX  
PF 10-APR-1998; 98WO-US006954.  
XX  
PR 10-APR-1997; 97US-00835913.  
XX  
PA (GEM) GENETICS INST INC.  
XX  
XX Jacobs K, McCooy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
PI Spaulding V, Agostino MJ;  
XX WPI; 1999-070076/06.  
XX  
XX New polynucleotides encoding human secreted proteins - derived from e.g.  
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
PT pituitary, retina and colon cDNA libraries.  
XX  
XX Claim 1; Page 131; 633pp; English.  
XX  
XX This sequence represents an expressed sequence tag (EST), and is a  
CC polynucleotide of the invention. The polynucleotides of the invention are

CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX SQ Sequence 396 BP; 100 A; 97 C; 101 G; 98 T; 0 U; 0 Other;

Query Match 92.0%; Score 18.4; DB 2; Length 396;  
 Best Local Similarity 95.0%; Pred. No. 19;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 ||||| ||||| ||||| ||||| |||||  
 DB 184 CAGACCCATAGACCAACAGG 203

RESULT 25  
 ADA02564/c  
 ID ADA02564 standard; DNA; 96596 BP.

XX AC ADA02564;

XX DT 06-NOV-2003 (first entry)

XX DE Human RASA2 carcinoma associated gene, SEQ ID NO:1082.

XX XW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 XW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 XW gene; ds.

XX OS Homo sapiens.

XX FN WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX PR 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-587068/55.

XX PT New recombinant nucleic acid encoding carcinoma associated protein,  
 PT useful for preparing compositions for treating carcinomas.

XX PS Claim 1; SEQ ID NO 1082; 245pp; English.

XX CC The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular

CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC sequence of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 96596 BP; 28171 A; 16866 C; 18487 G; 31716 T; 0 U; 1356 Other;

Query Match 87.0%; Score 17.4; DB 8; Length 96596;  
 Best Local Similarity 94.7%; Pred. No. 96;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAG 19  
 ||||| ||||| ||||| ||||| |||||  
 DB 69970 CAGACATAGACCAACAG 69952

RESULT 26  
 ADB72302/c  
 ID ADB72302 standard; DNA; 96596 BP.

XX AC ADB72302;

XX DT 04-DEC-2003 (first entry)

XX DR Human RASA2 gene.

XX XW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
 XW cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX OS Homo sapiens.

XX FN WO2003008583-A2.

XX PD 30-JAN-2003.

XX PF 26-DEC-2001; 2001WO-US051291.

XX PR 02-MAR-2001; 2001US-00798586.

XX PR 23-OCT-2001; 2001US-00004113.

XX PR 08-NOV-2001; 2001US-00052482.

XX PR 30-NOV-2001; 2001US-00997722.

XX PR 20-DEC-2001; 2001US-00034650.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW, Engelhard BK;

XX DR WPI; 2003-239337/23.

XX PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

XX PT cancers, neoplasm, adenocarcinoma, or sarcomas.

XX PS Claim 1; SEQ ID NO 130; 2304pp; English.

XX CC The invention relates to a novel recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the 660 sequences fully defined  
 CC in the specification. A polynucleotide of the invention has cytostatic  
 CC activity, and may have a use in gene therapy, or in a vaccine. The  
 CC recombinant nucleic acids and polypeptides are useful for treating  
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
 CC sarcomas. The present sequence represents a human gene of the invention.

XX SQ Sequence 96596 BP; 28171 A; 16866 C; 18487 G; 31716 T; 0 U; 1356 Other;

Query Match 87.0%; Score 17.4; DB 9; Length 96596;  
 Best Local Similarity 94.7%; Pred. No. 96;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAG 19  
 ||||| ||||| ||||| ||||| |||||

Db 69970 CAGACGATAGACCAACAG 69952

RESULT 27

ACD13446/C  
ID ACD13446 standard; DNA; 240000 BP.

XX ACD13446;

XX 13-AUG-2003 (first entry)

XX Human DNA encoding a p53 modifier, SEQ ID 31.

XX Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;  
XX antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;  
XX kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;  
XX apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX W0200299122-Al.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-US017382.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-156859/15.

XX P-PSDB; ABO07186.

XX Identifying modulators of the p53 pathway for use in treating apoptotic  
XX or cell proliferation disorders, comprises screening for agents that  
XX modulate activity of a human ortholog of genes that modify the p53  
XX pathway in Drosophila.

XX Example 2; Page 194-258; 678pp; English.

XX The invention relates to identifying (M1) a candidate p53 pathway  
XX modulating agent, by contacting an assay system comprising a purified HM  
XX polypeptide (human orthologue of genes that modify the p53 pathway in  
XX Drosophila) or nucleic acid with a test agent under conditions, where but  
XX for the presence of the test agent, the system provides a reference  
XX activity, and detecting a test agent-biased activity of the assay system.  
XX Also included are modulating (M2) a p53 pathway of a cell (comprising  
XX contacting a cell defective in p53 function with a candidate modulator  
XX that specifically binds to a HM polypeptide comprising an HM amino acid  
XX sequence, where p53 function is restored), modulating (M3) a p53 pathway  
XX in a mammalian cell (comprising contacting the cell with an agent that  
XX specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)  
XX a disease in a patient (comprising: (a) obtaining a biological sample  
XX from the patient; (b) contacting the sample with a probe for HM  
XX expression; (c) comparing the results with a control; and (d) determining  
XX whether the comparison indicates a likelihood disease). (M1) is useful  
XX for identifying modulators of the p53 pathway. A probe for HM expression  
XX is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,  
XX in a patient, where the cancer has greater than 25% expression level.  
XX Modulators identified by (M1) are useful in a variety of diagnostic and  
XX therapeutic applications, where disease or disorder prognosis is related  
XX to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell  
XX proliferation disorders (e.g. cancer). Another two new methods (M2 and  
XX M3) are useful for modulating the p53 pathway of a cell, thus restoring  
XX the p53 function of the cell, so that the cell undergoes normal  
XX proliferation or progression through the cell cycle. (M2) and (M3) are  
XX also useful for treating defects in the p53 pathway such as angiogenic,  
XX apoptotic or cell proliferation disorders. The present sequence is an HM  
XX nucleic acid encoding a p53 pathway modifying protein

XX Sequence 240000 B2; 74584 A; 44724 C; 45218 G; 75474 T; 0 U; 0 Other;  
SQ

Query Match 87.0%; Score 17.4; DB 7; Length 240000;  
Best Local Similarity 94.7%; Pred. No. 1e-02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGATAGACCAACAG 19

Db 176348 CAGACATAGACCAACAG 176330

RESULT 28

AAI99682\_03/c  
Continuation (4 of 45) of AAI99682 from base 300001 (Mycobacterium tuberculosis strain H37Rv)  
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682

WP	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000
WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 84.0%; Score 15.8; DB 4; Length 110000;  
Best Local Similarity 90.0%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGATAGACCAACAG 20

Db 81828 CAGACGATAGACCAACAG 81809

RESULT 29

AAI99683\_03/c  
Continuation (4 of 44) of AA-99683 from base 300001 (Mycobacterium tuberculosis strain H  
WP Sequence split into 44 fragments LOCUS AAI99683 Accession: AAI99683  
Fragment Name Begin End  
WP AAI99683\_00 1 110000  
WP AAI99683\_01 100001 210000  
WP AAI99683\_02 200001 310000  
WP AAI99683\_03 300001 410000  
WP AAI99683\_04 400001 510000  
WP AAI99683\_05 500001 610000  
WP AAI99683\_06 600001 710000  
WP AAI99683\_07 700001 810000  
WP AAI99683\_08 800001 910000  
WP AAI99683\_09 900001 1010000  
WP AAI99683\_10 1000001 1110000  
WP AAI99683\_11 1100001 1210000  
WP AAI99683\_12 1200001 1310000  
WP AAI99683\_13 1300001 1410000  
WP AAI99683\_14 1400001 1510000  
WP AAI99683\_15 1500001 1610000  
WP AAI99683\_16 1600001 1710000  
WP AAI99683\_17 1700001 1810000  
WP AAI99683\_18 1800001 1910000  
WP AAI99683\_19 1900001 2010000  
WP AAI99683\_20 2000001 2110000  
WP AAI99683\_21 2100001 2210000  
WP AAI99683\_22 2200001 2310000  
WP AAI99683\_23 2300001 2410000  
WP AAI99683\_24 2400001 2510000  
WP AAI99683\_25 2500001 2610000  
WP AAI99683\_26 2600001 2710000  
WP AAI99683\_27 2700001 2810000  
WP AAI99683\_28 2800001 2910000  
WP AAI99683\_29 2900001 3010000  
WP AAI99683\_30 3000001 3110000  
WP AAI99683\_31 3100001 3210000  
WP AAI99683\_32 3200001 3310000  
WP AAI99683\_33 3300001 3410000  
WP AAI99683\_34 3400001 3510000  
WP AAI99683\_35 3500001 3610000  
WP AAI99683\_36 3600001 3710000  
WP AAI99683\_37 3700001 3810000  
WP AAI99683\_38 3800001 3910000  
WP AAI99683\_39 3900001 4010000  
WP AAI99683\_40 4000001 4110000  
WP AAI99683\_41 4100001 4210000  
WP AAI99683\_42 4200001 4310000  
WP AAI99683\_43 4300001 4403765

Query Match 84.0%; Score 16.8; DB 4; Length 110000;  
Best Local Similarity 90.0%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGCGCATGACCAACAGG 20

Db 81884 CAGCGCATGACCAACAGG 81865

RESULT 30

AAZ49334/c

ID AAZ49334 standard; cDNA; 4189 BP.

XX AAZ49334;

AC AAZ49334;

XX AAZ49334;

DT 14-MAR-2000 (first entry)

DE Murine multidrug resistance-1 (MDR-1) cDNA.

XX Multidrug resistance; MDR-1; P-glycoprotein; transmembrane efflux pump;

KW haematopoietic stem cell; transduction; bone marrow transplantation;

KW chemotherapy; radiation therapy; cancer; gene therapy; gene replacement;

KW genetic defect; thalassaemia; Gaucher's disease; sickle cell anaemia;

KW leukaemia; ex vivo expansion; cytokine; ds.

XX Mus sp.  
OS  
FH Key Location/Qualifiers  
FT CDS 1..3831  
FT /\*tag= a  
FT /product= "Murine MDR-1 protein"  
XX WO9961589-A2.  
XX 02-DEC-1999.  
XX 27-MAY-1999; 99WO-US011825.  
XX 28-MAY-1998; 98JS-0086988P.  
XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX Sorrentino E, Bunting K;  
XX WPI; 2000-072615/06.  
XX P-ESDB; AAY58188.  
XX Ex vivo expansion of hematopoietic stem cells transduced with a sequence  
encoding human multidrug resistance-1, used for bone marrow  
transplantation.  
XX Disclosure; Page 90-93; 113pp; English.

This sequence represents cDNA encoding murine multidrug resistance  
protein MDR-1. MDR-1 is a transmembrane efflux pump, responsible for the  
export of drugs from cells, particularly cancer cells. The invention  
relates to transducing haematopoietic stem cells with nucleic acid  
encoding an MDR protein and culturing the modified cells. The modified  
haematopoietic stem cells are useful in bone marrow transplantation (to  
reconstitute haematopoietic systems in patients who have undergone  
chemotherapy or radiation therapy) and in ex vivo gene therapy of genetic  
defects in cells derived from haematopoietic stem cells, e.g.,  
thalassaemia, Gaucher's disease, sickle cell anaemia or leukaemia. The  
modified cells can also be used to identify factors involved in  
regulating proliferation and differentiation in haematopoietic stem  
cells. Haematopoietic stem cells that express MDR-1 will be protected  
against chemotherapeutic agents, so can be engrafted while the patient is  
undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
provides sufficient cells to permit standard biochemical analysis.  
Overexpression of MDR-1 allows cytokine-driven expansion of  
haematopoietic stem cells by at least 10-fold compared with a maximum of  
4-fold in known procedures

XX Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 3; Length 4189;

Best Local Similarity 94.4%; Pred. No. 2.3e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GAGCGATAGACCAACAGG 20

Db 923 GATGATAGACCAACAGG 906

RESULT 31

ABA94367/c

ID ABA94367 standard; DNA; 4189 BP.

XX ABA94367;

AC ABA94367;

XX ABA94367;

DT 26-MAR-2002 (first entry)

DE Mouse BCRP DNA related seq Id No. 5.

XX Stem cell; ATP transport protein; ATP-binding cassette; anti-parkinsonian;

KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;

KW cardiant; gene therapy; ds.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 ; Search time 27.1765 Seconds  
(without alignments)  
408.405 Million cell updates/sec

Title: US-09-877-819B-39

Perfect score: 20

Sequence: 1 cagacgcataagaccacagg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/ptctus\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	19	95.0	4233	3	US-09-120-513-1
C 2	19	95.0	4233	3	US-09-450-105-1
C 3	16.8	84.0	1860	4	US-09-894-844-65
C 4	16.8	84.0	4403765	3	US-09-103-840A-2
C 5	16.8	84.0	4411529	3	US-09-103-840A-1
C 6	16	80.0	148567	4	US-09-801-876B-3
C 7	16	80.0	148567	4	US-10-254-869-3
C 8	15.8	79.0	112132	4	US-09-741-150-3
C 9	15.8	79.0	112132	4	US-10-160-187-3
C 10	15.4	77.0	1068	4	US-09-328-352-2430
C 11	15.2	76.0	1178	3	US-08-861-774E-93
C 12	15.2	76.0	7785	2	US-08-276-967-1
C 13	14.8	74.0	483	4	US-09-483-039A-5676
C 14	14.8	74.0	1332	4	US-09-711-164-143
C 15	14.8	74.0	1335	4	US-09-480-039A-3751
C 16	14.8	74.0	3088	1	US-08-418-444A-1
C 17	14.8	74.0	20986	4	US-08-961-521-54
C 18	14.4	72.0	525	4	US-09-621-976-17799
C 19	14.4	72.0	894	4	US-09-107-532A-1317
C 20	14.4	72.0	1200	3	US-09-018-584A-37
C 21	14.4	72.0	1233	4	US-09-252-991A-2121
C 22	14.4	72.0	1395	4	US-09-252-991A-1954
C 23	14.4	72.0	3001	4	US-09-539-333D-167
C 24	14.4	72.0	5304	4	US-09-023-655-664
C 25	14.4	72.0	22067	4	US-09-820-001-3
C 26	14.4	72.0	50500	4	US-09-146-053-3
C 27	14.4	72.0	80246	3	US-09-078-294-4

C 28	14.4	72.0	80595	3	US-09-078-294-3	Sequence 3, Appli
C 29	14.4	72.0	128779	4	US-09-497-855A-38	Sequence 38, Appl
C 30	14.4	72.0	202001	4	US-09-734-674-3	Sequence 3, Appli
C 31	14.4	72.0	786431	4	US-09-751-389-3	Sequence 3, Appli
C 32	14.2	71.0	271	4	US-09-313-294A-33	Sequence 33, Appl
C 33	14.2	71.0	465	4	US-09-252-991A-2817	Sequence 2817, Ap
C 34	14.2	71.0	944	3	US-09-386-493-4	Sequence 4, Appli
C 35	14.2	71.0	1001	4	US-09-671-317-241	Sequence 241, App
C 36	14.2	71.0	1001	4	US-09-671-317-242	Sequence 242, App
C 37	14.2	71.0	1001	4	US-09-671-317-243	Sequence 243, App
C 38	14.2	71.0	1001	4	US-09-671-317-244	Sequence 244, App
C 39	14.2	71.0	1001	4	US-09-671-317-245	Sequence 245, App
C 40	14.2	71.0	1045	2	US-09-014-969-6	Sequence 6, Appli
C 41	14.2	71.0	1218	4	US-08-943-607-17	Sequence 17, Appl
C 42	14.2	71.0	1638	4	US-09-489-039A-3879	Sequence 3879, Ap
C 43	14.2	71.0	1713	3	US-09-347-798-7	Sequence 7, Appli
C 44	14.2	71.0	3621	4	US-09-252-991A-3061	Sequence 3061, Ap
C 45	14.2	71.0	3729	4	US-09-107-532A-1587	Sequence 1587, Ap
C 46	14.2	71.0	3902	4	US-09-976-594-792	Sequence 792, App
C 47	14.2	71.0	3988	4	US-09-762-195-1	Sequence 1, Appli
C 48	14.2	71.0	4066	4	US-09-367-750-1	Sequence 1, Appli
C 49	14.2	71.0	4264	2	US-08-784-649A-1	Sequence 1, Appli
C 50	14.2	71.0	4264	2	US-08-784-649A-5	Sequence 5, Appli
C 51	14.2	71.0	4646	1	US-08-181-471-2	Sequence 2, Appli
C 52	14.2	71.0	4646	1	US-09-023-655-1167	Sequence 1167, Ap
C 53	14.2	71.0	4669	2	US-08-583-276-18	Sequence 1, Appli
C 54	14.2	71.0	4669	2	US-08-752-447-1	Sequence 1, Appli
C 55	14.2	71.0	4669	4	US-09-316-167-1	Sequence 1, Appli
C 56	14.2	71.0	4669	4	US-09-397-233-1	Sequence 1, Appli
C 57	14.2	71.0	4669	6	5206352-3	Patent No. 5206352
C 58	14.2	71.0	5125	1	US-08-094-948A-4	Sequence 4, Appli
C 59	14.2	71.0	5125	5	PT-US96-09319-4	Sequence 4, Appli
C 60	14.2	71.0	6505	2	US-08-793-610-5	Sequence 5, Appli
C 61	14.2	71.0	8630	4	US-09-306-417-1	Sequence 1, Appli
C 62	14.2	71.0	8630	4	US-09-306-417-2	Sequence 2, Appli
C 63	14.2	71.0	9318	2	US-08-793-610-6	Sequence 6, Appli
C 64	14.2	71.0	12734	4	US-09-344-456-1	Sequence 1, Appli
C 65	14.2	71.0	14078	3	US-09-433-262-1	Sequence 1, Appli
C 66	14.2	71.0	14078	4	US-09-702-330-1	Sequence 1, Appli
C 67	14.2	71.0	14578	3	US-08-659-694-1	Sequence 1, Appli
C 68	14.2	71.0	55216	4	US-09-116-865-23	Sequence 23, Appl
C 69	13.8	69.0	293	4	US-09-313-294A-3503	Sequence 3503, Ap
C 70	13.8	69.0	481	4	US-09-621-976-66	Sequence 66, Appl
C 71	13.8	69.0	700	4	US-09-702-705-184	Sequence 184, App
C 72	13.8	69.0	700	4	US-09-736-457-184	Sequence 184, App
C 73	13.8	69.0	700	4	US-08-614-124B-184	Sequence 184, App
C 74	13.8	69.0	700	4	US-09-671-325-184	Sequence 184, App
C 75	13.8	69.0	900	3	US-09-589-184-184	Sequence 184, App
C 76	13.8	69.0	900	3	US-08-987-121A-1	Sequence 1, Appli
C 77	13.8	69.0	900	3	US-08-987-121A-2	Sequence 2, Appli
C 78	13.8	69.0	1008	3	US-08-987-121A-5	Sequence 5, Appli
C 79	13.8	69.0	1011	3	US-08-987-121A-3	Sequence 3, Appli
C 80	13.8	69.0	1011	3	US-09-066-512-1	Sequence 1, Appli
C 81	13.8	69.0	1095	4	US-09-328-352-533	Sequence 533, App
C 82	13.8	69.0	1446	4	US-09-489-039A-5151	Sequence 5151, Ap
C 83	13.8	69.0	1884	4	US-09-687-538B-5	Sequence 5, Appli
C 84	13.8	69.0	2007	3	US-08-941-445A-8	Sequence 8, Appli
C 85	13.8	69.0	2085	1	US-08-572-951-2	Sequence 2, Appli
C 86	13.8	69.0	4065	6	5240838-14	Patent No. 5240838
C 87	13.8	69.0	4379	1	US-08-592-214A-17	Sequence 17, Appl
C 88	13.8	69.0	4379	3	US-08-149-976-17	Sequence 17, Appl
C 89	13.8	69.0	36138	4	US-08-311-731A-136	Sequence 136, App
C 90	13.8	69.0	37030	4	US-08-311-731A-25	Sequence 25, Appl
C 91	13.8	69.0	43795	3	US-08-742-185-101	Sequence 101, App
C 92	13.8	69.0	193303	4	US-08-497-855A-37	Sequence 37, Appl
C 93	13.8	69.0	193303	4	US-09-497-855A-44	Sequence 44, Appl
C 94	13.6	68.0	88	4	US-09-465-355-29	Sequence 29, Appl
C 95	13.6	68.0	348	4	US-08-651-155B-255	Sequence 255, App
C 96	13.6	68.0	348	4	US-09-194-036B-255	Sequence 255, App
C 97	13.6	68.0	378	4	US-09-199-637A-316	Sequence 316, App
C 98	13.6	68.0	393	4	US-09-313-294A-4256	Sequence 4256, Ap
C 99	13.6	68.0	402	4	US-09-252-991A-7024	Sequence 7024, Ap
C 100	13.6	68.0	442	4	US-09-280-116-164	Sequence 164, App

101 13.6 68.0 513 4 US-09-621-976-613 Sequence 613, App  
c 102 13.6 68.0 576 4 US-09-252-991A-10362 Sequence 10362, A  
c 103 13.6 68.0 590 1 US-08-466-033-116 Sequence 116, App  
c 104 13.6 68.0 590 1 US-08-444-733-116 Sequence 116, App  
c 105 13.6 68.0 590 2 US-08-464-134-116 Sequence 116, App  
c 106 13.6 68.0 590 2 US-08-461-361-116 Sequence 116, App  
c 107 13.6 68.0 590 2 US-08-485-910-116 Sequence 116, App  
c 108 13.6 68.0 596 3 US-09-328-111-576 Sequence 576, App  
c 109 13.6 68.0 598 3 US-08-998-416-1090 Sequence 1090, App  
c 110 13.6 68.0 686 3 US-08-998-416-1040 Sequence 1040, App  
c 111 13.6 68.0 720 1 US-08-044-621D-1 Sequence 1, Appli  
c 112 13.6 68.0 741 4 US-09-489-039A-849 Sequence 849, App  
c 113 13.6 68.0 732 4 US-09-252-991A-7724 Sequence 7724, App  
c 114 13.6 68.0 819 4 US-09-199-637A-310 Sequence 310, App  
c 115 13.6 68.0 831 4 US-09-205-258-55 Sequence 55, Appl  
c 116 13.6 68.0 990 4 US-09-205-258-55 Sequence 55, Appl  
c 117 13.6 68.0 1014 4 US-09-199-637A-314 Sequence 314, App  
c 118 13.6 68.0 1014 4 US-09-252-991A-4829 Sequence 4829, App  
c 119 13.6 68.0 1026 4 US-09-540-236-1082 Sequence 1082, App  
c 120 13.6 68.0 1041 4 US-09-199-637A-322 Sequence 322, App  
c 121 13.6 68.0 1041 4 US-09-252-991A-4861 Sequence 4861, App  
c 122 13.6 68.0 1230 4 US-09-252-991A-7646 Sequence 7646, App  
c 123 13.6 68.0 1242 4 US-09-134-001C-544 Sequence 544, App  
c 124 13.6 68.0 1302 3 US-08-332-823A-1 Sequence 1, Appli  
c 125 13.6 68.0 1349 4 US-09-570-856B-2 Sequence 2, Appli  
c 126 13.6 68.0 1384 4 US-09-205-258-225 Sequence 225, App  
c 127 13.6 68.0 1459 4 US-09-180-109A-13 Sequence 13, Appli  
c 128 13.6 68.0 1472 4 US-09-540-224-3 Sequence 3, Appli  
c 129 13.6 68.0 1472 4 US-09-564-595D-52 Sequence 52, Appli  
c 130 13.6 68.0 1472 4 US-09-808-972-3 Sequence 3, Appli  
c 131 13.6 68.0 1497 4 US-09-252-991A-10612 Sequence 10612, A  
c 132 13.6 68.0 1704 4 US-09-252-991A-4855 Sequence 4855, App  
c 133 13.6 68.0 1744 4 US-09-221-017B-887 Sequence 887, App  
c 134 13.6 68.0 2268 4 US-09-620-312D-909 Sequence 909, App  
c 135 13.6 68.0 2690 1 US-08-524-757-11 Sequence 11, Appl  
c 136 13.6 68.0 2731 4 US-09-016-434-1344 Sequence 1344, App  
c 137 13.6 68.0 4242 4 US-09-252-991A-7056 Sequence 7056, App  
c 138 13.6 68.0 4665 4 US-09-252-991A-10243 Sequence 10243, A  
c 139 13.6 68.0 4843 4 US-09-060-299-41 Sequence 41, Appl  
c 140 13.6 68.0 4843 4 US-09-402-923A-41 Sequence 41, Appl  
c 141 13.6 68.0 5117 4 US-09-060-299-40 Sequence 40, Appl  
c 142 13.6 68.0 5117 4 US-09-402-923A-40 Sequence 40, Appl  
c 143 13.6 68.0 7286 4 US-08-793-273C-1 Sequence 1, Appli  
c 144 13.6 68.0 7286 5 PCT-US95-1168A-1 Sequence 1, Appli  
c 145 13.6 68.0 7724 4 US-08-466-049-1 Sequence 1, Appli  
c 146 13.6 68.0 9103 1 US-08-466-033-182 Sequence 182, App  
c 147 13.6 68.0 9103 1 US-08-444-733-182 Sequence 182, App  
c 148 13.6 68.0 9103 2 US-08-464-134-182 Sequence 182, App  
c 149 13.6 68.0 9103 2 US-08-461-361-182 Sequence 182, App  
c 150 13.6 68.0 9103 2 US-08-485-910-182 Sequence 182, App  
c 151 13.6 68.0 9103 5 PCT-US95-06266-156 Sequence 156, App  
c 152 13.6 68.0 9327 1 US-08-466-033-234 Sequence 234, App  
c 153 13.6 68.0 9327 1 US-08-444-733-234 Sequence 234, App  
c 154 13.6 68.0 9327 2 US-08-464-134-234 Sequence 234, App  
c 155 13.6 68.0 9327 2 US-08-461-361-234 Sequence 234, App  
c 156 13.6 68.0 9327 2 US-08-485-910-234 Sequence 234, App  
c 157 13.6 68.0 9391 1 US-08-638-911A-1 Sequence 1, Appli  
c 158 13.6 68.0 9391 5 PCT-US95-06266-14 Sequence 14, Appli  
c 159 13.6 68.0 9392 1 US-08-466-033-14 Sequence 14, Appli  
c 160 13.6 68.0 9392 1 US-08-444-733-14 Sequence 14, Appli  
c 161 13.6 68.0 9392 2 US-08-464-134-14 Sequence 14, Appli  
c 162 13.6 68.0 9392 2 US-08-461-361-14 Sequence 14, Appli  
c 163 13.6 68.0 9392 2 US-08-485-910-14 Sequence 14, Appli  
c 164 13.6 68.0 10023 4 US-09-252-991A-6997 Sequence 6997, App  
c 165 13.6 68.0 16075 3 US-09-096-942-1 Sequence 1, Appli  
c 166 13.6 68.0 16075 3 US-09-096-867-1 Sequence 1, Appli  
c 167 13.6 68.0 19988 4 US-09-596-002-10 Sequence 10, Appli  
c 168 13.6 68.0 2108 3 US-09-053-197A-3 Sequence 3, Appli  
c 169 13.6 68.0 2108 4 US-09-085-761A-3 Sequence 3, Appli  
c 170 13.6 68.0 31328 4 US-09-215-694-19 Sequence 19, Appli  
c 171 13.6 68.0 36241 4 US-08-311-731A-134 Sequence 134, App  
c 172 13.6 68.0 36470 4 US-08-311-731A-123 Sequence 123, App  
c 173 13.6 68.0 36941 4 US-08-311-731A-130 Sequence 130, App

174 13.6 68.0 42235 4 US-09-199-637A-1 Sequence 1, Appli  
175 13.6 68.0 45175 4 US-09-453-702B-116 Sequence 116, App  
176 13.6 68.0 56109 4 US-09-598-002-35 Sequence 35, Appl  
c 177 13.6 68.0 56165 4 US-09-214-808-1 Sequence 1, Appli  
c 178 13.6 68.0 786431 4 US-09-751-389-3 Sequence 3, Appli  
179 13.6 68.0 4403765 3 US-09-103-840A-2 Sequence 2, Appli  
180 13.6 68.0 4411529 3 US-09-103-840A-1 Sequence 1, Appli

US-09-120-513-1/c  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Ellens, Harma  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4233 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
US-09-120-513-1

Query Match 95.0%; Score 19; DB 3; Length 4233;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGACGCATAGACCAACAGG 20  
Db 946 AGACGCATAGACCAACAGG 928

RESULT 2  
US-09-450-105-1/c  
; Sequence 1, Application US/09450105  
; Patent No. 6169166  
; GENERAL INFORMATION:



APPLICANT: Kimberly Anne Brun  
APPLICANT: Richard James Chenery  
APPLICANT: Hama Ellens  
APPLICANT: John Anthony Feild  
APPLICANT: Lin Yue  
TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
FILE REFERENCE: GP-50008-D1  
CURRENT FILING DATE: 1999-11-29  
EARLIER APPLICATION NUMBER: 09/120,513  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4233  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
US-09-450-105-1

Query Match 95.0%; Score 19; DB 3; Length 4233;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAGG 20  
DB 946 AGACGCATAGACCAACAGG 928  
|||||

RESULT 3  
US-09-894-844-65/c  
Sequence 65, Application US/09894844  
Patent No. 6686166  
GENERAL INFORMATION:  
APPLICANT: Behr, Marcel  
APPLICANT: Small, Peter  
APPLICANT: Schoolnik, Gary  
APPLICANT: Wilson, Michael A.  
TITLE OF INVENTION: Molecular Differences Between Species of  
TITLE OF INVENTION: the M. Tuberculosis Complex  
FILE REFERENCE: STAN102CON  
CURRENT APPLICATION NUMBER: US/09/894,844  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: 09/318,191  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/097,936  
PRIOR FILING DATE: 1998-08-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 65  
LENGTH: 1860  
TYPE: DNA  
ORGANISM: Mycobacteria tuberculosis  
US-09-894-844-65

Query Match 84.0%; Score 16.8; DB 4; Length 1860;  
Best Local Similarity 90.0%; Pred. No. 1.3;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 1275 CAGACGCATAGACCAACAGG 1256  
|||||

RESULT 4  
US-09-103-840A-2/c  
Sequence 2, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 84.0%; Score 16.8; DB 3; Length 4403765;  
Best Local Similarity 90.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 381884 CAGACGCATAGACCAACAGG 381865  
|||||

RESULT 5  
US-09-103-840A-1/c  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 84.0%; Score 16.8; DB 3; Length 4411529;  
Best Local Similarity 90.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 381828 CAGACGCATAGACCAACAGG 381809  
|||||

RESULT 6  
US-09-801-876B-3/c  
Sequence 3, Application US/09801876B  
Patent No. 6492155  
GENERAL INFORMATION:  
APPLICANT: YE, Jane et al  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CLO01160  
CURRENT APPLICATION NUMBER: US/09/801,876B  
CURRENT FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3

```
/ LENGTH: 148567
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURES:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(148567)
/ OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3

Query Match      80.0%; Score 16; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAAA 16
Db 141057 CAGACGCATAGACCAAA 141042

RESULT 7
US-10-254-869-3/c
/ Sequence 3, Application US/10254869
/ Patent No. 6653117
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: CL001160DIV
/ CURRENT APPLICATION NUMBER: US/10/254.869
/ CURRENT FILING DATE: 2002-09-26
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 148567
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(148567)
/ OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

Query Match      80.0%; Score 16; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAAA 16
Db 141057 CAGACGCATAGACCAAA 141042

RESULT 8
US-09-741-150-3
/ Sequence 3, Application US/09741150
/ Patent No. 6436689
/ GENERAL INFORMATION:
/ APPLICANT: GUEGLER, Karl et al
/ TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: CL000968
/ CURRENT APPLICATION NUMBER: US/09/741,150
/ CURRENT FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 112132
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(112132)
/ OTHER INFORMATION: n = A,T,C or G

US-09-741-150-3
Query Match      79.0%; Score 15.8; DB 4; Length 112132;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAG 19
Db 72903 CAGACGCATAGACCAATAG 72921

RESULT 9
US-10-160-187-3
/ Sequence 3, Application US/10160187
/ Patent No. 6620607
/ GENERAL INFORMATION:
/ APPLICANT: GUEGLER, Karl et al.
/ TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: CL000968DIV
/ CURRENT APPLICATION NUMBER: US/10/160,187
/ CURRENT FILING DATE: 2002-06-04
/ PRIOR APPLICATION NUMBER: 60/252,410
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: 09/741,150
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 112132
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(112132)
/ OTHER INFORMATION: n = A,T,C or G
US-10-160-187-3

Query Match      79.0%; Score 15.8; DB 4; Length 112132;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAG 19
Db 72903 CAGACGCATAGACCAATAG 72921

RESULT 10
US-09-328-352-2430/c
/ Sequence 2430, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 2430
/ LENGTH: 1068
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
/ OTHER INFORMATION:
US-09-328-352-2430

Query Match      77.0%; Score 15.4; DB 4; Length 1068;
Best Local Similarity 94.1%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CAGCGATAGACCAACAG 19
Db 670 GACGCATAGACCAACGG 654
```

```
RESULT 11
US-08-861-774B-93
; Sequence 93, Application US/08861774E
; Patent No. 6297007
; GENERAL INFORMATION:
; APPLICANT: Waters, Barbara
; APPLICANT: Miao, Vivian
; APPLICANT: Ho, Yap
; APPLICANT: Tong, Seow
; TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR
; FILE REFERENCE: 9993-006
; CURRENT APPLICATION NUMBER: US/08/861,774E
; CURRENT FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 93
; LENGTH: 1178
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Clone ps3
US-08-861-774E-93

Query Match      76.0%; Score 15.2; DB 3; Length 1178;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
DB 726 CGGACGATAGGACCAACAGG 745

RESULT 12
US-08-276-967-1/c
; Sequence 1, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:418\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7785 base pairs
; TYPE: nucleic acid
```

```
STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-276-967-1

Query Match      76.0%; Score 15.2; DB 2; Length 7785;
Best Local Similarity 85.0%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
DB 550 CAGAAGCATAGACCATGAGG 531

RESULT 13
US-09-489-039A-5676
; Sequence 5676, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gazy Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5676
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5676

Query Match      74.0%; Score 14.8; DB 4; Length 483;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GACGCATAGACCAACAGG 20
DB 355 GAAGCATATACCAACAGG 372

RESULT 14
US-09-711-164-143/c
; Sequence 143, Application US/09711164
; Patent No. 6889738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERET
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1332)
US-09-711-164-143

Query Match      74.0%; Score 14.8; DB 4; Length 1332;
Best Local Similarity 88.9%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACA 18
```

```
Db          970 CAGACGCATAGACCAACA 953
|||||
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-418-444A-1

Query Match          74.0%; Score 14.8; DB 1; Length 3088;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          1 CAGACGCATAGACCAACA 18
|||||
Db          2978 CAGACCCATAGATCAACA 2961
|||||

RESULT 17
US-08-961-527-54/c
; Sequence 54, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-54

Query Match          74.0%; Score 14.8; DB 4; Length 20986;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          3 CAGCGCATAGACCAACAGG 20
|||||
Db          4421 GTCCCATAGACCAAGAGG 4404
|||||

RESULT 18
US-09-621-976-17799
; Sequence 17799, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
```

```
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.Pm
; SEQ ID NO 17799
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17799

Query Match      72.0%; Score 14.4; DB 4; Length 525;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAA 16
      |||||
Db      25 CAGACGCATAGACCAA 40

RESULT 19
US-09-107-532A-1317/c
; Sequence 1317, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051,571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1317:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (B) LOCATION 1...894
; SEQUENCE DESCRIPTION: SEQ ID NO: 1317:
US-09-107-532A-1317

Query Match      72.0%; Score 14.4; DB 4; Length 894;
Best Local Similarity 93.8%; Pred. No. 2.e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      5 CGCATAGACCAACAGG 20
      |||||
Db      265 CGCTAGACCAACAGG 250

RESULT 20
US-09-018-584A-37
; Sequence 37, Application US/09018584A
; Patent No. 6238863
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; APPLICANT: Bachter, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
; REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,584A
; FILING DATE: 04-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1200 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: no
; IMMEDIATE SOURCE:
; CLONE: S199
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 6q21
US-09-018-584A-37

Query Match      72.0%; Score 14.4; DB 3; Length 1200;
Best Local Similarity 93.8%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAA 16
      |||||
Db      1020 CAGACACATAGACCAA 1035

RESULT 21
US-09-252-991A-2-21/c
; Sequence 2121, Application US/09252991A
```

; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2121  
; LENGTH: 1233  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-2121

Query Match 72.0%; Score 14.4; DB 4; Length 1233;  
Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCATAGACCAACAGG 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1093 CCCATAGACCAACAGG 1078

RESULT 22  
US-09-252-991A-1954/c  
; Sequence 1954, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1954  
; LENGTH: 1395  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1954

Query Match 72.0%; Score 14.4; DB 4; Length 1395;  
Best Local Similarity 93.8%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCATAGACCAACAGG 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1296 CCCATAGACCAACAGG 1281

RESULT 23  
US-09-539-333D-167/c  
; Sequence 167, Application US/09539333D  
; Patent No. 6476208  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bouqueloret, Lydie  
; APPLICANT: Bihain, Bernard  
; APPLICANT: Esieux, Laurent  
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS  
; FILE REFERENCE: GENSET.047AUS  
; CURRENT APPLICATION NUMBER: US/09/539,333D

; CURRENT FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: US 60/126,903  
; PRIOR FILING DATE: 1999-03-30  
; PRIOR APPLICATION NUMBER: US 60/131,971  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: US 60/132,065  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: US 60/143,928  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 60/145,915  
; PRIOR FILING DATE: 1999-07-27  
; PRIOR APPLICATION NUMBER: US 60/146,453  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: US 60/146,452  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: US 60/162,288  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: US 09/416,384  
; PRIOR FILING DATE: 1999-10-12  
; NUMBER OF SEQ ID NOS: 231  
; SOFTWARE: Patent.pm  
; SEQ ID NO 167  
; LENGTH: 3001  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 1501  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 1502\_1521  
; OTHER INFORMATION: 99-26167-278.misl, complement  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 1482..1500  
; OTHER INFORMATION: 99-26167-278.misl  
; FEATURE:  
; NAME/KEY: primer\_bind  
; LOCATION: 1759..1778  
; OTHER INFORMATION: upstream amplification primer, complement  
; FEATURE:  
; NAME/KEY: primer\_bind  
; LOCATION: 1319..1339  
; OTHER INFORMATION: downstream amplification primer  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 1489..1513  
; OTHER INFORMATION: 99-26167-278 probe  
US-09-539-333D-167

Query Match 72.0%; Score 14.4; DB 4; Length 3001;  
Best Local Similarity 93.8%; Pred. No. 2.4e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGACATAGACCAA 16  
| | | | | | | | | | | | | | | | | |  
Db 981 CAGACGACATAGACCAA 966

RESULT 24  
US-09-023-655-664/c  
; Sequence 664, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	20	100.0	20	10	US-09-877-819B-39	Sequence 39, Appl
2	20	100.0	21	10	US-09-877-819B-40	Sequence 40, Appl
C	3	20	100.0	254	US-09-877-819B-55	Sequence 55, Appl
4	20	100.0	267	9	US-09-796-632-143	Sequence 143, App
5	20	100.0	267	9	US-09-796-632-328	Sequence 328, App
6	20	100.0	267	9	US-09-796-632-7075	Sequence 7075, App
7	20	100.0	267	15	US-10-040-862-143	Sequence 143, App
8	20	100.0	267	15	US-10-040-862-328	Sequence 328, App
9	20	100.0	267	15	US-10-040-862-7075	Sequence 7075, App
10	20	100.0	267	16	US-10-057-475B-143	Sequence 143, App
11	20	100.0	267	16	US-10-057-475B-328	Sequence 328, App
12	20	100.0	267	16	US-10-057-475B-7075	Sequence 7075, App
13	20	100.0	267	16	US-10-154-884B-143	Sequence 143, App
14	20	100.0	267	16	US-10-154-884B-328	Sequence 328, App

88	15.8	79.0	564	16	US-10-027-632-35201	Sequence 35201, A	C 161	15.2	76.0	77478	13	US-10-087-192-1378	Sequence 1378, Ap
89	15.8	79.0	564	16	US-10-027-632-35202	Sequence 35202, A	162	15.2	75.0	3309400	9	US-09-738-626-1	Sequence 1, Appl
90	15.8	79.0	582	13	US-10-027-632-238862	Sequence 238862, A	C 163	15.2	75.0	1014	16	US-10-369-493-40938	Sequence 40938, A
91	15.8	79.0	582	13	US-10-027-632-238863	Sequence 238863, A	164	14.8	74.0	128	13	US-10-424-599-92120	Sequence 92120, A
92	15.8	79.0	582	13	US-10-027-632-238864	Sequence 238864, A	165	14.8	74.0	191	13	US-10-282-122A-2522	Sequence 2522, Ap
93	15.8	79.0	582	13	US-10-027-632-238865	Sequence 238865, A	166	14.8	74.0	223	13	US-10-282-122A-2900	Sequence 2900, Ap
94	15.8	79.0	582	16	US-10-027-632-238866	Sequence 238866, A	167	14.8	74.0	229	13	US-10-424-599-4467	Sequence 4467, Ap
95	15.8	79.0	582	16	US-10-027-632-238867	Sequence 238867, A	168	14.8	74.0	239	13	US-10-282-122A-2630	Sequence 2630, Ap
96	15.8	79.0	585	13	US-10-027-632-79299	Sequence 79299, A	169	14.8	74.0	239	13	US-10-282-122A-2716	Sequence 2716, Ap
97	15.8	79.0	585	16	US-10-027-632-79299	Sequence 79299, A	170	14.8	74.0	248	13	US-10-282-122A-2541	Sequence 2541, Ap
98	15.8	79.0	611	13	US-10-027-632-220215	Sequence 220215, A	171	14.8	74.0	278	13	US-10-282-122A-2745	Sequence 2745, Ap
99	15.8	79.0	611	13	US-10-027-632-220216	Sequence 220216, A	172	14.8	74.0	303	13	US-10-282-122A-2598	Sequence 2598, Ap
100	15.8	79.0	611	13	US-10-027-632-220217	Sequence 220217, A	173	14.8	74.0	325	13	US-10-282-122A-2594	Sequence 2594, Ap
101	15.8	79.0	611	16	US-10-027-632-220215	Sequence 220215, A	174	14.8	74.0	354	13	US-10-282-122A-2742	Sequence 2742, Ap
102	15.8	79.0	611	16	US-10-027-632-220216	Sequence 220216, A	175	14.8	74.0	355	13	US-10-282-122A-2594	Sequence 2594, Ap
103	15.8	79.0	615	13	US-10-027-632-24892	Sequence 24892, A	176	14.8	74.0	455	9	US-09-783-590-5421	Sequence 5421, Ap
104	15.8	79.0	615	16	US-10-027-632-24892	Sequence 24892, A	177	14.8	74.0	456	13	US-10-282-122A-2370	Sequence 2370, Ap
105	15.8	79.0	615	16	US-10-027-632-214892	Sequence 214892, A	178	14.8	74.0	457	13	US-10-282-122A-3057	Sequence 3057, Ap
106	15.8	79.0	1329	13	US-10-424-599-6159	Sequence 6159, Ap	179	14.8	74.0	468	13	US-10-027-632-214522	Sequence 214522, A
107	15.8	79.0	108382	13	US-10-087-192-1618	Sequence 1618, Ap	180	14.8	74.0	468	16	US-10-027-632-214522	Sequence 214522, A
108	15.8	79.0	194049	13	US-10-087-192-1495	Sequence 1495, Ap							
109	15.8	79.0	302603	13	US-10-271-416-8	Sequence 8, Appl							
110	15.8	79.0	322101	13	US-10-354-247-1	Sequence 1, Appl							
111	15.8	79.0	322101	15	US-10-060-902-1	Sequence 1, Appl							
112	15.4	77.0	465	13	US-10-085-783A-57491	Sequence 57491, A							
113	15.4	77.0	465	16	US-10-242-535A-57491	Sequence 57491, A							
114	15.4	77.0	1257	13	US-10-027-632-124090	Sequence 124090, A							
115	15.4	77.0	1257	16	US-10-027-632-124090	Sequence 12409							



OTHER INFORMATION: Address sequence  
US-09-877-819B-40

Query Match 100.0%; Score 20; DB 10; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
Db 21 CAGACGCATAGACCAACAGG 2

RESULT 3  
US-09-877-819B-55  
Sequence 55, Application US/09877819B  
Publication No. US20030190609A1  
GENERAL INFORMATION:  
APPLICANT: Torrey, David  
TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing  
FILE REFERENCE: S-94,664  
CURRENT APPLICATION NUMBER: US/09/877,819B  
CURRENT FILING DATE: 2001-06-07  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 55  
LENGTH: 254  
TYPE: DNA  
ORGANISM: Human HLA  
US-09-877-819B-55

Query Match 100.0%; Score 20; DB 10; Length 254;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
Db 40 CAGACGCATAGACCAACAGG 59

RESULT 4  
US-09-796-692-143  
Sequence 143, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 328  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-796-692-328

Query Match 100.0%; Score 20; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
Db 161 CAGACGCATAGACCAACAGG 180

PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 143  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-796-692-143

Query Match 100.0%; Score 20; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 5  
US-09-796-692-328  
Sequence 328, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 328  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-796-692-328

```

RESULT 6
US-09-796-692-7075
; Sequence 7075, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7075

Query Match      100.0%; Score 20; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CAGACGCATAGACCAACAGG 20
   |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 7
US-10-040-862-143
; Sequence 143, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01

```

```

; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-143

Query Match      100.0%; Score 20; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CAGACGCATAGACCAACAGG 20
   |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 8
US-10-040-862-328
; Sequence 328, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; CURRENT APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950

```

```
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-328
```

```
Query Match 100.0%; Score 20; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
    ||||||||||||||||
Db 161 CAGACGCATAGACCAACAGG 180
```

```
RESULT 9
US-10-040-862-7075
; Sequence 7075, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7075
```

US-10-040-862-7075

```
Query Match 100.0%; Score 20; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
    ||||||||||||||||
Db 161 CAGACGCATAGACCAACAGG 180
```

```
RESULT 10
US-10-057-475B-143
; Sequence 143, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-143
```

```
Query Match 100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
    ||||||||||||||||
Db 161 CAGACGCATAGACCAACAGG 180
```

```
RESULT 11
US-10-057-475B-328
; Sequence 328, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-328

Query Match 100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 12
US-10-057-475B-7075
; Sequence 7075, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
```

```
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-7075

Query Match 100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 161 CAGACGCATAGACCAACAGG 180

RESULT 13
US-10-154-884B-143
; Sequence 143, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-143

Query Match          100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
DB      161 CAGACGCATAGACCAACAGG 180

RESULT 14
US-10-154-884B-328
; Sequence 328, Application US/10154884B
; Publication No. US2004000561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-7075

Query Match          100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
DB      161 CAGACGCATAGACCAACAGG 180

RESULT 16
US-09-796-692-42
; Sequence 42, Application US/05796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-328

Query Match          100.0%; Score 20; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
DB      161 CAGACGCATAGACCAACAGG 180

RESULT 15
US-10-154-884B-7075
; Sequence 7075, Application US/10154884B
; Publication No. US2004000561A1
```

US-09-796-692-3554

Query Match 100.0%; Score 20; DB 9; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 166 CAGACGCATAGACCAACAGG 185

RESULT 17

US-09-796-692-3554

Sequence 3554, Application US/09796692  
Publication No. US20020198362A1

GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 42

LENGTH: 272

TYPE: DNA

ORGANISM: Homo sapiens

US-09-796-692-42

Query Match 100.0%; Score 20; DB 9; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 166 CAGACGCATAGACCAACAGG 185

RESULT 18

US-10-040-862-42

Sequence 42, Application US/10040862  
Publication No. US20030078396A1

GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/040,862

PRIOR FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: US 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 42

LENGTH: 272

TYPE: DNA

ORGANISM: Homo sapiens

US-09-796-692-3554

Query Match 100.0%; Score 20; DB 9; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 166 CAGACGCATAGACCAACAGG 185

RESULT 19

US-10-040-862-3554

Sequence 3554, Application US/10040862  
Publication No. US20030078396A1

GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3554

LENGTH: 272

TYPE: DNA

ORGANISM: Homo sapiens

US-10-040-862-42

Query Match 100.0%; Score 20; DB 15; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
DB 166 CAGACGCATAGACCAACAGG 185

RESULT 19

US-10-040-862-3554

Sequence 3554, Application US/10040862  
Publication No. US20030078396A1

GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/040,862

PRIOR FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 09/796,692

PRIOR FILING DATE: 2001-03-01

NUMBER OF SEQ ID NOS: 10467

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 42

LENGTH: 272

TYPE: DNA

ORGANISM: Homo sapiens

```
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/040,862
/ CURRENT FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3554
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-040-862-3554

Query Match      100.0%; Score 20; DB 15; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
Db      166 CAGACGCATAGACCAACAGG 185

RESULT 20
US-10-057-475B-42
/ Sequence 42, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordonez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/040,862
/ CURRENT FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 42
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-42

Query Match      100.0%; Score 20; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
Db      166 CAGACGCATAGACCAACAGG 185

RESULT 21
US-10-057-475B-3554
/ Sequence 3554, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordonez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
```

```
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-3554
```

```
Query Match 100.0%; Score 20; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 166 CAGACGCATAGACCAACAGG 185
```

```
RESULT 22
US-10-154-884B-42
; Sequence 42, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,125
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-42
```

```
Query Match 100.0%; Score 20; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 166 CAGACGCATAGACCAACAGG 185
```

```
RESULT 23
US-10-154-884B-3554
; Sequence 3554, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-3554
```

```
Query Match 100.0%; Score 20; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
|||||
Db 166 CAGACGCATAGACCAACAGG 185
```

```
RESULT 24
US-09-796-692-280/c
; Sequence 280, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 2002-03-01
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```



```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-280
```

```
Query Match 100.0%; Score 20; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CAGACGCATAGACCAACAGG 20
|||
Db 107 CAGACGCATAGACCAACAGG 88
```

```
RESULT 25
US-09-796-692-530/c
; Sequence 530, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
```

```
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-530
```

```
Query Match 100.0%; Score 20; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CAGACGCATAGACCAACAGG 20
|||
Db 107 CAGACGCATAGACCAACAGG 88
```

```
RESULT 26
US-09-796-692-3270/c
; Sequence 3270, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-3270
```

```
Query Match 100.0%; Score 20; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CAGACGCATAGACCAACAGG 20
|||
Db 107 CAGACGCATAGACCAACAGG 88
```

```
RESULT 27
US-10-040-862-280/c
```

```
; Sequence 280, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-280
```

```
Query Match 100.0%; Score 20; DB 15; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CAGACGCATAGACCAACAGG 20
Db 107 CAGACGCATAGACCAACAGG 88
```

```
RESULT 28
US-10-040-862-530/c
; Sequence 530, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
```

```
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-530
```

```
Query Match 100.0%; Score 20; DB 15; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CAGACGCATAGACCAACAGG 20
Db 107 CAGACGCATAGACCAACAGG 88
```

```
RESULT 29
US-10-040-862-3270/c
; Sequence 3270, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
```

;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3270  
;; LENGTH: 294  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-3270

Query Match 100.0%; Score 20; DB 15; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 107 CAGACGCATAGACCAACAGG 88

RESULT 30  
US-10-057-475B-280/c  
;; Sequence 280, Application US/100574753  
;; Publication No. US20040002068A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Clapper, Jonathan David  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Ordenez, Nadia  
;; APPLICANT: Carter, Lauren  
;; APPLICANT: McNeill, Patricia Dianne  
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy of Hematological Malignancies  
;; FILE REFERENCE: 014058-014402US  
;; CURRENT APPLICATION NUMBER: US 60/1057,475B  
;; CURRENT FILING DATE: 2002-01-22  
;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 10979  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 280  
;; LENGTH: 294  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-057-475B-280

Query Match 100.0%; Score 20; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 107 CAGACGCATAGACCAACAGG 88

RESULT 31  
US-10-057-475B-530/c  
;; Sequence 530, Application US/10057475B  
;; Publication No. US20040002068A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Clapper, Jonathan David  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Ordenez, Nadia  
;; APPLICANT: Carter, Lauren  
;; APPLICANT: McNeill, Patricia Dianne  
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy of Hematological Malignancies  
;; FILE REFERENCE: 014058-014402US  
;; CURRENT APPLICATION NUMBER: US 60/1057,475B  
;; CURRENT FILING DATE: 2002-01-22  
;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 10979  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 530  
;; LENGTH: 294  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-057-475B-530

Query Match 100.0%; Score 20; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 107 CAGACGCATAGACCAACAGG 88

RESULT 32  
US-10-057-475B-3270/c  
;; Sequence 3270, Application US/10057475B  
;; Publication No. US20040002068A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander

```
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US 10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-3270

Query Match      100.0%; Score 20; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
      |||||
Db      107 CAGACGCATAGACCAACAGG 88

RESULT 33
US-10-154-884B-280/c
; Sequence 280, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US 10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-280

Query Match      100.0%; Score 20; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
      |||||
Db      107 CAGACGCATAGACCAACAGG 88

RESULT 34
US-10-154-884B-530/c
; Sequence 530, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US 10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-10-154-884B-530
Query Match      100.0%; Score 20; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
Db      107 CAGACGCATAGACCAACAGG 88

RESULT 35
US-10-154-884B-3270/c
; Sequence 3270, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-3270

Query Match      100.0%; Score 20; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
Db      107 CAGACGCATAGACCAACAGG 88

RESULT 36
US-09-918-995-36460
; Sequence 36460, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
```

```
US-09-918-995-36460
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36460
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-36460

Query Match      100.0%; Score 20; DB 10; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
Db      184 CAGACGCATAGACCAACAGG 203

RESULT 37
US-09-918-995-37140
; Sequence 37140, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37140
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-37140

Query Match      100.0%; Score 20; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      |||
Db      257 CAGACGCATAGACCAACAGG 276

RESULT 38
US-10-102-524-715
; Sequence 715, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 715
```

```
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-715

Query Match      100.0%; Score 20; DB 15; Length 523;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
Db      172 CAGACGCATAGACCAACAGG 191

RESULT 39
US-10-084-817-88
; Sequence 88, Application US/10094817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shoret
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 88
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1090035.1
US-10-084-817-88

Query Match      100.0%; Score 20; DB 15; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
Db      205 CAGACGCATAGACCAACAGG 224

RESULT 40
US-10-220-120-138
; Sequence 138, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.

; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKDREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;
60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
60/184,763; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 138
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:1169865.1:2000MAY01
US-10-220-120-138

Query Match      100.0%; Score 20; DB 13; Length 1259;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
Db      347 CAGACGCATAGACCAACAGG 366

RESULT 41
US-09-925-302-351
; Sequence 351, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 1348
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1294)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1318)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1329)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-351
```

```

Query Match      100.0%; Score 20; DB 9; Length 1348;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CAGACGCATAGACCAACAGG 20
Db      285 CAGACGCATAGACCAACAGG 304
```

```

RESULT 42
US-09-925-302-351
; Sequence 351, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1294)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1318)
; OTHER INFORMATION: n equals a,t,g, or c
```

```

; NAME/KEY: misc feature
; LOCATION: (1329)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-351
```

```

Query Match      100.0%; Score 20; DB 13; Length 1348;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 CAGACGCATAGACCAACAGG 20
Db      285 CAGACGCATAGACCAACAGG 304
```

```

RESULT 43
US-09-960-706-1043
; Sequence 1043, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)...(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-960-706-1043
```

```

Query Match      100.0%; Score 20; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 CAGACGCATAGACCAACAGG 20
Db      6366 CAGACGCATAGACCAACAGG 6385
```

```

RESULT 44
US-09-873-319-691
; Sequence 891, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 691
; LENGTH: 14646
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X03100
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-873-319-691

Query Match          100.0%; Score 20; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 6366 CAGACGCATAGACCAACAGG 6385

RESULT 45
US-09-917-800A-1424/c
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M91855
US-09-917-800A-1424

Query Match          95.0%; Score 19; DB 9; Length 4254;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAGG 20
    |||||
Db 1026 AGACGCATAGACCAACAGG 1008

RESULT 46
US-10-388-934-265/c
; Sequence 265, Application US/10388934
```

```
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 265
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-265

Query Match          95.0%; Score 19; DB 16; Length 4254;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAGG 20
    |||||
Db 1026 AGACGCATAGACCAACAGG 1008

RESULT 47
US-10-027-632-46603/c
; Sequence 46603, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46603
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46603

Query Match          87.0%; Score 17.4; DB 13; Length 696;
Best Local Similarity 94.7%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAG 19
    |||||
Db 236 CAGACGCATAGACCAACAG 218

RESULT 48
```



```
US-10-027-632-46603/c
; Sequence 46603, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46603
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46603

Query Match      87.0%; Score 17.4; DB 16; Length 696;
Best Local Similarity 94.7%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAG 19
Db 236 CAGACATAGACCAACAG 218

RESULT 49
US-10-027-632-22275/c
; Sequence 22275, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22275
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-22275
```

```
US-10-027-632-22276/c
; Sequence 22276, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22276
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-22276

Query Match      84.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAG 20
Db 333 CAGACATAGACCAACAG 314

RESULT 50
US-10-027-632-22277/c
; Sequence 22277, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22277
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-22277

Query Match      84.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAG 20
Db 333 CAGACATAGACCAACAG 314

RESULT 51
US-10-027-632-22277/c
; Sequence 22277, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22277
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-22277
```

```
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222277
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222277
```

```
Query Match 84.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
||||| ||||| ||||| |||||
Db 333 CAGACACATAGACCAAAAGG 314
```

## RESULT 52

```
US-10-027-632-222278/c
; Sequence 222278, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222278
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222278
```

```
Query Match 84.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
||||| ||||| ||||| |||||
Db 333 CAGACACATAGACCAAAAGG 314
```

## RESULT 53

```
US-10-027-632-222275/c
; Sequence 222275, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222275
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222275
```

```
Query Match 84.0%; Score 16.8; DB 16; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
||||| ||||| ||||| |||||
Db 333 CAGACACATAGACCAAAAGG 314
```

## RESULT 54

```
US-10-027-632-222276/c
; Sequence 222276, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222276
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222276
```

```
Query Match 84.0%; Score 16.8; DB 16; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CAGACGCATAGACCAACAGG 20
||||| ||||| ||||| |||||
Db 333 CAGACACATAGACCAAAAGG 314
```

```
RESULT 55
US-10-027-632-222277/c
; Sequence 222277, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222277
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222277

Query Match      84.0%; Score 16.8; DB 16; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      ||||| ||||| ||||| |||||
Db      333 CAGACACATAGACCAAAAGG 314

RESULT 56
US-10-027-632-222278/c
; Sequence 222278, Application US/10027632
; Publication No. US20030204075A5
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222278
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
```

```
; ORGANISM: Human
US-10-027-632-222278

Query Match      84.0%; Score 16.8; DB 16; Length 572;
Best Local Similarity 90.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      ||||| ||||| ||||| |||||
Db      333 CAGACACATAGACCAAAAGG 314

RESULT 57
US-09-894-844-65/c
; Sequence 65, Application US/09894844
; Patent No. US20020176873A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-65

Query Match      84.0%; Score 16.8; DB 9; Length 1860;
Best Local Similarity 90.0%; Pred. No. 86;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
      ||||| ||||| ||||| |||||
Db      1275 CAGACGCATCGACACAGG 1256

RESULT 58
US-10-647-089-65/c
; Sequence 65, Application US/10647089
; Publication No. US20040063923A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/10/647,089
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
```

US-10-647-089-65

Query Match 84.0%; Score 16.8; DB 13; Length 1860;  
Best Local Similarity 90.0%; Pred. No. 86;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
|||||||  
DB 1275 CAGACGCATCGACCAACAGG 1256

RESULT 59

US-10-388-902-65/c

; Sequence 65, Application US/10388902  
; Publication No. US20040018574A1  
; GENERAL INFORMATION:  
; APPLICANT: Behr, Marcel  
; APPLICANT: Small, Peter  
; APPLICANT: Schoolnik, Gary  
; APPLICANT: Wilson, Michael A.  
; TITLE OF INVENTION: Molecular Differences Between Species of  
; FILE OF INVENTION: the M. Tuberculosis Complex

; CURRENT APPLICATION NUMBER: US/10/388,902  
; CURRENT FILING DATE: 2003-03-14  
; PRIOR APPLICATION NUMBER: US/09/894,844  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/318,191  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 60/097,936  
; PRIOR FILING DATE: 1998-08-25  
; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 65

; LENGTH: 1860

; TYPE: DNA

; ORGANISM: Mycobacteria tuberculosis

US-10-388-902-65

Query Match 84.0%; Score 16.8; DB 16; Length 1860;  
Best Local Similarity 90.0%; Pred. No. 86;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
|||||||  
DB 1275 CAGACGCATCGACCAACAGG 1256

RESULT 60

US-10-425-114-34249/c

; Sequence 34249, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 34249

; LENGTH: 1256

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLMO17236F01\_FLI

US-10-425-114-34249

Query Match 82.0%; Score 16.4; DB 13; Length 1256;  
Best Local Similarity 94.4%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACGCATAGACCAACAG 19  
|||||||  
DB 883 AGACGCATAGACCAACAG 866

Search completed: April 20, 2004, 12:56:14

Job time : 126.706 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:56:28 ; Search time 1272.82 Seconds  
(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819B-39

Perfect score: 20

Sequence: 1 cagacgcatagaccacag 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	147	10	BE140679
2	20	100.0	159	14	T91602
C 3	20	100.0	229	10	BE140710
4	20	100.0	230	13	C03623

78	20	100.0	720	10	BE976100	BF976100	602244936	151	20	100.0	1055	14	CD517144	CD517144	AGENCOURT
79	20	100.0	724	10	BE439689	BE439689	HTML-360P	152	20	100.0	1061	12	BQ054682	BQ054682	AGENCOURT
80	20	100.0	732	12	BG599921	BG599921	602681285	153	20	100.0	1083	13	BQ594971	BQ594971	AGENCOURT
81	20	100.0	732	14	CB997879	CB997879	AGENCOURT	154	20	100.0	1108	13	BQ072806	BQ072806	AGENCOURT
82	20	100.0	734	9	AV733676	AV733676		155	20	100.0	1111	14	CD517173	CD517173	AGENCOURT
83	20	100.0	735	12	BI762063	BI762063	603049235	156	20	100.0	1201	13	EX363746	EX363746	AGENCOURT
84	20	100.0	743	14	CB956955	CB956955	AGENCOURT	157	20	100.0	1217	12	BM544066	BM544066	AGENCOURT
85	20	100.0	743	14	CD104670	CD104670	AGENCOURT	158	20	100.0	1722	13	BQ057412	BQ057412	AGENCOURT
86	20	100.0	744	12	BI597401	BI597401	603243611	159	19	95.0	338	9	AA323639	AA323639	EST126729
87	20	100.0	745	14	CB984378	CB984378	AGENCOURT	160	18.4	92.0	177	10	BG025808	BG025808	602274621
88	20	100.0	748	10	BF129007	BF129007	601811329	161	18.4	92.0	256	10	BE159476	BE159476	MRO-HT040
89	20	100.0	751	12	BG759996	BG759996	602733419	162	18.4	92.0	334	14	CD704658	CD704658	EST21185
90	20	100.0	758	14	CB985816	CB985816	AGENCOURT	163	18.4	92.0	405	9	AA244273	AA244273	nc060606.r
91	20	100.0	761	12	BG541682	BG541682	602571259	164	18.4	92.0	408	10	BM834672	BM834672	K-EST0109
92	20	100.0	766	12	BG570800	BG570800	602591286	165	18.4	92.0	439	10	AM406315	AM406315	UI-HF-PL0
93	20	100.0	776	12	BI227278	BI227278	602948666	166	18.4	92.0	443	14	CD687150	CD687150	EST3671.h
94	20	100.0	789	12	BG538956	BG538956	602568315	167	18.4	92.0	444	14	CD704753	CD704753	CD704753
95	20	100.0	790	12	BI837330	BI837330	603090674	168	18.4	92.0	448	10	BF834707	BF834707	CM3-HT096
96	20	100.0	791	13	BG596174	BG596174	AGENCOURT	169	18.4	92.0	464	14	CD705070	CD705070	EST21597
97	20	100.0	796	12	BG399870	BG399870	602441984	170	18.4	92.0	483	14	CD692032	CD692032	EST8555.h
98	20	100.0	796	14	CB997069	CB997069	AGENCOURT	171	18.4	92.0	494	14	CD698716	CD698716	EST15239
99	20	100.0	802	10	BF238242	BF238242	601811922	172	18.4	92.0	511	14	CD707609	CD707609	EST24136
100	20	100.0	802	12	BG757550	BG757550	602714723	173	18.4	92.0	527	14	CD684349	CD684349	EST869.m
101	20	100.0	806	14	CB993192	CB993192	AGENCOURT	174	18.4	92.0	530	14	CD708787	CD708787	EST25314
102	20	100.0	809	14	CB992720	CB992720	AGENCOURT	175	18.4	92.0	533	14	CD687511	CD687511	EST4032.h
103	20	100.0	811	12	BI766898	BI766898	603053138	176	18.4	92.0	534	10	BG756165	BG756165	602713466
104	20	100.0	821	12	BI668794	BI668794	603294742	177	18.4	92.0	546	10	AM351777	AM351777	RC3-CT019
105	20	100.0	823	14	CB991431	CB991431	AGENCOURT	178	18.4	92.0	548	14	CD696409	CD696409	EST12932
106	20	100.0	826	12	BG431523	BG431523	602498354	179	18.4	92.0	588	14	CD698528	CD698528	EST15051
107	20	100.0	830	14	CB994002	CB994002	AGENCOURT	180	18.4	92.0	590	14	CD707303	CD707303	EST23830
108	20	100.0	831	10	BF525771	BF525771	602069914								
109	20	100.0	832	14	CB958376	CB958376	AGENCOURT								
110	20	100.0	837	12	BI517599	BI517599	603042205								
111	20	100.0	843	12	BI261592	BI261592	602953662								
112	20	100.0	851	14	CD558780	CD558780	AGENCOURT								
113	20	100.0	854	12	BG756171	BG756171	602713472								
114	20	100.0	857	13	BQ890384	BQ890384	AGENCOURT								
115	20	100.0	859	12	BG546165	BG546165	602573582								
116	20	100.0	860	12	BG755507	BG755507	602713882								
117	20	100.0	871	12	BI767351	BI767351	603057883								
118	20	100.0	872	12	BG542978	BG542978	602569375								
119	20	100.0	873	12	BQ721555	BQ721555	AGENCOURT								
120	20	100.0	873	12	BG540219	BG540219	602569108								
121	20	100.0	878	14	CB986657	CB986657	AGENCOURT								
122	20	100.0	880	9	AU139061	AU139061									
123	20	100.0	881	14	CD558622	CD558622	AGENCOURT								
124	20	100.0	884	12	BI545349	BI545349	603187481								
125	20	100.0	889	14	CD516536	CD516536	AGENCOURT								
126	20	100.0	890	12	BG754449	BG754449	602710081								
127	20	100.0	901	13	BQ955105	BQ955105	AGENCOURT								
128	20	100.0	911	10	BE878872	BE878872	601493167								
129	20	100.0	915	10	BF974675	BF974675	602245314								
130	20	100.0	917	12	BG397584	BG397584	602438625								
131	20	100.0	923	13	BQ642494	BQ642494	AGENCOURT								
132	20	100.0	926	13	BQ642196	BQ642196	AGENCOURT								
133	20	100.0	943	13	BQ660004	BQ660004	AGENCOURT								
134	20	100.0	945	12	BG536146	BG536146	602564242								
135	20	100.0	953	10	BF975606	BF975606	602246064								
136	20	100.0	959	10	BF976257	BF976257	602245136								
137	20	100.0	959	13	BQ057748	BQ057748	AGENCOURT								
138	20	100.0	960	13	BQ894372	BQ894372	AGENCOURT								
139	20	100.0	967	12	BI553587	BI553587	603190506								
140	20	100.0	980	10	BF974318	BF974318	602243975								
141	20	100.0	980	13	BQ056978	BQ056978	AGENCOURT								
142	20	100.0	980	12	BG586860	BG586860	602550843								
143	20	100.0	992	12	BG759351	BG759351	602711866								
144	20	100.0	1005	13	BQ062257	BQ062257	AGENCOURT								
145	20	100.0	1007	13	BQ057757	BQ057757	AGENCOURT								
146	20	100.0	1015	12	BQ060930	BQ060930	AGENCOURT								
147	20	100.0	1025	12	BQ054502	BQ054502	AGENCOURT								
148	20	100.0	1026	13	BQ061344	BQ061344	AGENCOURT								
149	20	100.0	1027	13	BQ064032	BQ064032	AGENCOURT								
150	20	100.0	1030	13	BQ063085	BQ063085	AGENCOURT								

## ALIGNMENTS

RESULT 1  
BE140679/c

BE140679 147 bp mRNA linear EST 21-JUN-2000  
CMO-HT0016-170695-004 HT0016 Homo sapiens cDNA, mRNA sequence.

BE140679.1 GI:8603400

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

EST.

FEATURES  
source

Location/Qualifiers  
1. .147  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0016"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (J.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 147;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAGCGCATAGACCAACAGG 20  
|||||  
Db 134 CAGAGCGCATAGACCAACAGG 115  
|||||

## RESULT 2

T91602 159 bp mRNA linear EST 22-MAR-1995  
LOCUS  
DEFINITION  
IMAGE:118349 5' similar to gb:K01506 HLA CLASS II  
HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN);, mRNA  
sequence.

ACCESSION  
VERSION  
T91602.1 GI:723515  
EST.

KEYWORDS  
SOURCE  
Homo sapiens (human)

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

AUTHORS  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, X., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Roeding, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE  
Generation and analysis of 280,000 human expressed sequence tags

JOURNAL  
Genome Res. 6 (9), 807-828 (1996)

MEDLINE  
97044478

PUBMED  
889549

COMMENT  
Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 114

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: M13RPI

High quality sequence stop: 114.

Location/Qualifiers

1. .159

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:496638"

/db\_xref="taxon:9606"

/clone="IMAGE:118349"

/sex="male"

/dev\_stage="72 years"

/lab\_host="SOLR cells (kanamycin resistant)";

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 159;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAGCGCATAGACCAACAGG 20  
|||||  
Db 84 CAGAGCGCATAGACCAACAGG 103  
|||||

## RESULT 3

BE140710/c 229 bp mRNA linear EST 21-JUN-2000  
LOCUS  
DEFINITION  
CMO-HT0016-140699-008 HT0016 Homo sapiens cDNA, mRNA sequence.

ACCESSION  
BE140710

VERSION  
BE140710.1 GI:8603431

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 229)

## AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE  
Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

## COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2=CM0-HT0016-140

699-008&t3=1999-06-14&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 7

High quality sequence stop: 229.

## FEATURES

Location/Qualifiers

1. .229

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0016"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 229;

Best Local Similarity 100.0%; Pred. No. 41;

```

Matches      20; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1      CAGACGCATAGACCAACAGG 20
      |||||
      122 CAGACGCATAGACCAACAGG 103

RESULT 4
LOCUS      C03623      230 bp      mRNA      linear      EST 30-JUL-1996
DEFINITION      C03623 Human heart cDNA (YNakamura) Homo sapiens cDNA clone
ACCESSION      C03623
VERSION      C03623.1 GI:1466874
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
Nakamura,Y.
Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
96299762
8661126
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. .230
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="3NHCI920"
/dev_stage="adult"
/clone_lib="Human heart cDNA (YNakamura)"
/note="Organ: heart; normalized directionally cloned cDNA
from adult heart"

FEATURES
source
Query Match      100.0%; Score 20; DB 13; Length 230;
Best Local Similarity      100.0%; Pred.No. 41;
Matches      20; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1      CAGACGCATAGACCAACAGG 20
      |||||
      196 CAGACGCATAGACCAACAGG 215

RESULT 5
LOCUS      CD707370      249 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION      EST23897 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION      CD707370
VERSION      CD707370.1 GI:32238000
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
Transcriptional Gene Expression Profile of Human Nasopharynx
Unpublished (2003)
Contact: Yixin Zeng
Cancer Center

```

```

Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
Location/Qualifiers
1. .249
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 249;
Best Local Similarity      100.0%; Pred.No. 42;
Matches      20; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1      CAGACGCATAGACCAACAGG 20
      |||||
      217 CAGACGCATAGACCAACAGG 236

RESULT 6
LOCUS      AW799168      263 bp      mRNA      linear      EST 16-MAY-2000
DEFINITION      RC0-UM0051-010300-011-c05 UM0051 Homo sapiens cDNA, mRNA sequence.
ACCESSION      AW799168
VERSION      AW799168.1 GI:7851038
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 263)
Dias Neto,E., Garcia Correa,R., Vertovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W.Jr., Zago,X.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc.Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC0-UM0051-010
300-011-c05&t3=2000-03-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 29
High quality sequence stop: 263.
Location/Qualifiers
1. .263
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0051"
/note="Organ: uterus; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.

```



196,716 ~ Ludwig Institute for Cancer Research] profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 263;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAGCGCATAGACCAACAGG 20

|||||  
Db 29 CAGAGCGCATAGACCAACAGG 48

## RESULT 7

BE242285

LOCUS

DEFINITION TCAAP1E1562 Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1562, mRNA sequence.

## ACCESSION

BE242285

VERSION BE242285.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 265)

AUTHORS Wei, Y., Teang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R.,

Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project

Unpublished (2000)

CONTACT: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@ccc.org

Citation: Carninci, P. and Hayashizaki, Y. High efficiency

full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M13 primer.

FEATURES

Location/Qualifiers

1..265

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TCAAP1562"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="myeloid cell"

/dev\_stage="pediatric 6 years"

/lab\_host="DH10B"

/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project-TCAA"

/note="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGGAGGAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGAGTCGATCGCGCGCGCAATAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pSB vector. Library went through one round of

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Garninci P, Westover A, Nishiyama Y, Ohsumi T,

Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,

Schneider C, Hayashizaki Y, High efficiency selection of

full-length cDNA by improved biotinylated cap trapper.,

DNA Res 4: 1, 61-6, Feb 28, 1997)"

## ORIGIN

Query Match

Best Local Similarity

Matches 20; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Query Match 100.0%; Score 20; DB 10; Length 265;

Best Local Similarity 100.0%; Pred. No. 43;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAGCGCATAGACCAACAGG 20

|||||  
Db 188 CAGAGCGCATAGACCAACAGG 207

## RESULT 8

AA360953

LOCUS

DEFINITION

EST70157 T-cell lymphoma Homo sapiens cDNA 5' end similar to major histocompatibility complex, class II antigen, alpha chain (GB:X03100), mRNA sequence.

## ACCESSION

AA360953

VERSION AA360953.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 289)

AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,

Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,

White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,

Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,

Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,

Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,

Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,

Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,

Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,

Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,

He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,

Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,

Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,

Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,

Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THCI72266

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..289

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="ATCC (inhost):165076"

/db\_xref="taxon:9606"

/cell\_type="T-lymphocyte"

/clone\_lib="T-cell lymphoma"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:

XhoI"

## ORIGIN

Query Match

Best Local Similarity

Matches 20; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

```

QY      1 CAGAGCATAGACCAACAGG 20
Db      163 CAGAGCATAGACCAACAGG 182

RESULT 9
LOCUS   CD706205                290 bp mRNA linear EST 25-JUN-2003
DEFINITION EST22732 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD706205
VERSION   CD706205.1 GI:32236835
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 290)
AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
          Zeng,Y.-X.
          Transcriptional Gene Expression Profile of Human Nasopharynx
          Unpublished (2003)
          Contact: Yixin Zeng
          Cancer Center
          Sun Yat-sen University
          651 Dongfeng Road East, Guangzhou 510060, China
          Tel: 86-1380-9770-743
          Fax: 86-20-8775-4506
          Email: yxzeng@gzsums.edu.cn.
FEATURES             source
          Location/Qualifiers
            1..290
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /tissue_type="normal nasopharynx"
              /clone_lib="human nasopharynx"
              /note="ESTs generated from a normal nasopharynx cDNA
              library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 290;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAGCATAGACCAACAGG 20
Db      201 CAGAGCATAGACCAACAGG 220

RESULT 10
LOCUS   CD694567                295 bp mRNA linear EST 25-JUN-2003
DEFINITION EST11090 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD694567
VERSION   CD694567.1 GI:32219338
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
          Zeng,Y.-X.
          Transcriptional Gene Expression Profile of Human Nasopharynx
          Unpublished (2003)
          Contact: Yixin Zeng
          Cancer Center
          Sun Yat-sen University
          651 Dongfeng Road East, Guangzhou 510060, China
          Tel: 86-1380-9770-743
          Fax: 86-20-8775-4506
          Email: yxzeng@gzsums.edu.cn.
FEATURES             source
          Location/Qualifiers
            1..295

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 295;
Best Local Similarity 100.0%; Pred.No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAGCATAGACCAACAGG 20
Db      192 CAGAGCATAGACCAACAGG 211

RESULT 11
LOCUS   BM772902                335 bp mRNA linear EST 04-MAR-2002
DEFINITION X-EST0057161 S1SNUS Homo sapiens cDNA clone S1SNUS-8-C07 5', mRNA
sequence.
ACCESSION BM772902
VERSION   BM772902.1 GI:19102517
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 335)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
          Kim,Y.S.
          21C Frontier Korean EST Project 2001
          Unpublished (2002)
          Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 8 row: C column: 07
          High quality sequence stop: 335.
          Location/Qualifiers
            1..335
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="S1SNUS-8-C07"
              /sex="F"
              /tissue_type="Ascites"
              /cell_type="lymphoblast-like"
              /cell_line="SNU-5"
              /lab_host="Top10F"
              /clone_lib="S1SNUS"
              /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
              Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
              bacterial alkaline phosphatase (BAP) and then deapped
              with tabacco acid pyrophosphatase (TAP). The deapped
              intact mRNA was ligated with DNA-RNA linker including EcoR
              I site by treatment of T4 RNA ligase and the first strand
              cDNA was synthesized from oligo dt-selected mRNA by
              priming with dt-tailed vector. The dt-tailed vector was
              adjusted to have about 50nt. The cDNA vector was
              circularized with E. coli DNA ligase after digestion of
              EcoRI which site is also included in vector. An RNA strand
              converted to a DNA strand by Okayama-Berg method. The
              obtained cDNA vectors were used for transformation of
              competent cells E. coli Top10F by electroporation method.
              The cDNA libraries constructed by this method are
              full-length enriched cDNA library."

```

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20  
 Db 212 CAGACGCATAGACCAACAGG 231

## RESULT 12

T94759  
 LOCUS T94759 366 bp mRNA linear EST 24-MAR-1995  
 DEFINITION y337f12.r1 Stratagene lung (#937210) Homo sapiens cDNA clone  
 IMAGE:119951 5' similar to 9b:K01506 HLA CLASS II  
 HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN);, mRNA  
 sequence.

ACCESSION T94759  
 VERSION T94759.1 GI:728247

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
 Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,  
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,N.,  
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
 Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.,  
 and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
 JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

PUBMED 889549

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1152

High quality sequence stops: 251 Source: IMAGE Consortium, LML  
 This clone is available royalty-free through LML; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1152 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 251.

## FEATURES

source

Location/Qualifiers

1..386

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:498240"

/db\_xref="taxon:9606"

/clone="IMAGE:119951"

/sex="male"

/dev\_stage="72 years"

/lab\_host="SOLR cells (kanamycin resistant)"

/clone\_lib="Stratagene lung (#937210)"

/note="Organ: lung; Vector: pBluescript SK-; Site: 1:

ECORI, Site:2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. normal lung. Average insert size: 1.0 Kb;

Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGAGG

3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

## Db

82 CAGACGCATAGACCAACAGG 101

## RESULT 13

CB267061

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.

TITLE EST analysis of human adipose gene expression

JOURNAL Unpublished (2002)

COMMENT

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, H497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGAGCGCGCCATGTGTGTGGT

BACKWARD: AATAGACTACTATAGGCGAATTGG

Seq primer: GTTGGTACCGGGAATTC.

Location/Qualifiers

1..387

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/sex="Male and Female"

/tissue\_type="Adipose"

/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

/note="Vector: lambdaTriplex"

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 387;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

Db 163 CAGACGCATAGACCAACAGG 182

## RESULT 14

CD709023

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and

Zeng,Y.-X.

TITLE Transcriptional Gene Expression Profile of Human Nasopharynx

JOURNAL Unpublished (2003)

COMMENT

Contact: Yixin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

|||||

82 CAGACGCATAGACCAACAGG 101

CB267061 387 bp mRNA linear EST 20-FEB-2003  
 1005967 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 cDNA 5', mRNA sequence.

CB267061

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.

TITLE EST analysis of human adipose gene expression

JOURNAL Unpublished (2002)

COMMENT

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, H497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGAGCGCGCCATGTGTGTGGT

BACKWARD: AATAGACTACTATAGGCGAATTGG

Seq primer: GTTGGTACCGGGAATTC.

Location/Qualifiers

1..387

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/sex="Male and Female"

/tissue\_type="Adipose"

/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

/note="Vector: lambdaTriplex"

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 387;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

Db 163 CAGACGCATAGACCAACAGG 182

CD709023 390 bp mRNA linear EST 25-JUN-2003  
 EST25550 human nasopharynx Homo sapiens cDNA, mRNA sequence.

CD709023

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and

Zeng,Y.-X.

TITLE Transcriptional Gene Expression Profile of Human Nasopharynx

JOURNAL Unpublished (2003)

COMMENT

Contact: Yixin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

Email: yxzeng@zsums.edu.cn.  
 Location/Qualifiers  
 1..390  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN  
 Query Match 100.0%; Score 20; DB 14; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 221 CAGACGCATAGACCAACAGG 240  
 |||||

RESULT 15  
 AA838010  
 LOCUS  
 DEFINITION  
 ce89d07.s1 NCI CGAP Col12 Homo sapiens cDNA clone IMAGE:1418797  
 similar to gb:K01505 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)  
 ALPHA CHAIN (HUMAN); contains Alu repetitive element;; mRNA  
 sequence.

ACCESSION  
 AA838010  
 VERSION  
 EST.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 1 (bases 1 to 415)  
 Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.  
 Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing Genomics 35 (1), 231-235 (1996)

FEATURES  
 source  
 1..415  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="3NHC1710"  
 /dev\_stage="adult"  
 /clone\_lib="Human heart cDNA (Ynakamura)"  
 /note="Organ: heart; normalized directionally cloned cDNA from adult heart"

ORIGIN  
 Query Match 100.0%; Score 20; DB 13; Length 422;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20  
 |||||  
 Db 207 CAGACGCATAGACCAACAGG 226  
 |||||

RESULT 17  
 CD102141  
 LOCUS  
 DEFINITION  
 AGENCOURT 14007825 NIH MGC 186 Homo sapiens cDNA clone IMAGE:30370583 5', mRNA sequence.

ACCESSION  
 CD102141  
 VERSION  
 EST.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 1 (bases 1 to 423)  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapb-remail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: CLONTech Laboratories, Inc.

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: NDCM159 row: m column: 24  
 High quality sequence stop: 417.

## FEATURES

source

Location/Qualifiers

1..423  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30370583"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_186"  
 /notes="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfII  
 (ggccattatggcc); Site 2: SfII (ggccctcggcc); Library is  
 oligo-dt primed and directionally cloned. cDNA was  
 prepared from a pooled samples of tissues from Skin,  
 meninges, duramater, pia matter and choroid plexus.  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CAGGCGCATATGCGC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCGCGGCGGCGCATG-dt(30)BN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
 Library"

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

Db 223 CAGACGCATAGACCAACAGG 242

## RESULT 18

AW085969/c

LOCUS

DEFINITION xc763d02.x1 NCI\_CGAP\_Ov32 Homo sapiens cDNA clone IMAGE:2590179 3',  
 similar to 90:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)  
 ALPHA CHAIN (HUMAN);, mRNA sequence.

ACCESSION AW085969

VERSION AW085969.1 GI:5041175

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 427)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuquiqui,  
 M.D., Michael R. Emert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

DNA Library Arrayed by: The I.M.A.G.E. Consortium, LLNL

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -400P from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1..427

/organism="Homo sapiens"

## FEATURES

source

/mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2590179"  
 /sex="female"  
 /tissue\_type="papillary serous carcinoma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Ov32"  
 /note="Organ: ovary; Vector: pAMP1; mRNA made from ovarian  
 carcinoma, cDNA made by oligo-dt priming.  
 Non-directionally cloned. Size-selected on agarose gel,  
 average insert size 500 bp. Non-amplified library."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

Db 303 CAGACGCATAGACCAACAGG 284

## RESULT 19

AW449605/c

LOCUS

DEFINITION UT-H-BI3-aku-g-06-0-UL.s1 NCI\_CGAP\_Sub5 Homo sapiens cDNA clone  
 IMAGE:2735794 3', mRNA sequence.

ACCESSION AW449605

VERSION AW449605.1 GI:6990311

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 427)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

The sequence contained an oligo-dt track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:  
 NCI-CGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..427

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2735794"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI\_CGAP\_Sub5"

/note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; NCI\_CGAP\_Sub5  
 is a subtracted library derived from NCI\_CGAP\_Sub4. The  
 NCI\_CGAP\_Sub5 library had 3 million recombinants. A  
 single-stranded DNA preparation of NCI\_CGAP\_Sub4 was used  
 as a tracer in a subtractive hybridization with a driver  
 comprising: the IMAGE pool (NCI\_CGAP\_Kid3 pool 1 LLAM  
 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)  
 1322376-1323911, 1456008-1456775, 1500552-1502855);  
 NCI\_CGAP\_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778  
 (IMAGE Clones); NCI\_CGAP\_Lu5 pool 1 LLAM 3575-3582,  
 1492104-1493255); NCI\_CGAP\_Lu5 pool 1 LLAM 3575-3582,  
 3851-3854 (IMAGE Clones)  
 1414920-1417991, 1520904-1522439); NCI\_CGAP\_GC4 pool 1 LLAM  
 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones)

## FEATURES

source

1257096-1258631, 1469064-1470983, 1475592-1476743);  
 NCI CGAP\_P222 pool 1 LHAM 2457-2459, 2758-2759, 3062-3068  
 (IMAGE Clones 985608-986759, 1101192-1101959,  
 1217928-1220615); NCI CGAP\_Co10 pool 1 LHAM 2644-2653,  
 2871-2872 (IMAGE Clones 1057416-1061255, 1144584-1145351). (10% of the driver  
 population), plus a pool of 3,840 arrayed clones from  
 NCI CGAP\_Sub1 (IMAGE Clones 2708616-2710535) and  
 NCI CGAP\_Sub2 (IMAGE Clones 2710536-2712455) (10% of  
 the driver population), plus a pool of 11,136 clones from  
 NCI CGAP\_Sub3 (IMAGE Clones 2712456-2723591) (10% of the  
 driver population), plus a pool of 5,472 clones from  
 NCI CGAP\_Sub4 (IMAGE Clones 2723592-2728969) (70% of the  
 driver population). Subtraction was performed as  
 previously described (Bonaldo, Lennon & Soares (1996):  
 Normalization and Subtraction: Two Approaches to  
 Facilitate Gene Discovery. Genome Research 6, 791-806.  
 TAG TISSUE=kidney  
 TAG LIB=NCI CGAP\_Kids  
 TAG\_SEQ=ATTC"

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 417 CAGACGCATAGACCAACAGG 398

## RESULT 20

BI710946

DEFINITION lq95b07.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:5023429  
 5' similar to SW-HAQZ HUMAN P20036 H1A CLASS II HISTOCOMPATIBILITY  
 ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

## ACCESSION

BI710946

VERSION BI710946.1 GI:15686641

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS

1 (bases 1 to 429)  
 Melton,D., Brown,J., Kenty,G., Permatt,A., Lee,C., Kaestner,K.,  
 Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,  
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,  
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisshvili,R.,  
 Williams,T., Jackson,Y. and Bowers,Y.

## TITLE

Endocrine Pancreas Consortium

Unpublished (2000)

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
 University Genome Sequencing Center For information on obtaining a  
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 340.

Location/Qualifiers

1. 429

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5023429"

## FEATURES

source

/tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
 XhoI; Site\_2: EcoRI; Constructed with lambda ZAPII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/est/lambda\_protocol.shtml).  
 Please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permatt  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 211 CAGACGCATAGACCAACAGG 230

## RESULT 21

BM767805

LOCUS

DEFINITION K-EST0050302 S1SNUS2 Homo sapiens cDNA clone S1SNUS2-12-E02 5',

mRNA sequence.

VERSION BM767805

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 437)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 12 row: E column: 02

High quality sequence stop: 437.

Location/Qualifiers

1. 437

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S1SNUS2-12-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-5"

/lab\_host="Top10F"

/clone\_lib="S1SNUS2"

/note="Organ: Stomach; Vector: pCNS; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A) + RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10<sup>+</sup> with electroporation method."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

|||||

Db 213 CAGACGCATAGACCAACAGG 232

## RESULT 22

BM766631  
LOCUS K-EST0048574 S1SNU52 Homo sapiens cDNA clone S1SNU52-10-E02 5',  
DEFINITION mRNA sequence.

ACCESSION BM766631.1 GI:19096246  
VERSION  
KEYWORDS  
SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 469)

REFERENCE  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience &amp; Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 10 row: E column: 02

High quality sequence stop: 469.

Location/Qualifiers

1..469

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S1SNU52-10-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-5"

/lab\_host="Top10<sup>+</sup>"

/clone\_lib="S1SNU52"

/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped

FEATURES  
source

with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10<sup>+</sup> with electroporation method."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 469;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20

|||||

Db 213 CAGACGCATAGACCAACAGG 232

## RESULT 23

AV734557  
LOCUS AV734557 482 bp mRNA linear EST 17-OCT-2000  
DEFINITION AV734557 cda Homo sapiens cDNA clone cdABEB10 5', mRNA sequence.

ACCESSION AV734557

VERSION AV734557.1 GI:10852102

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 482)

REFERENCE  
AUTHORS Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,  
Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S.,  
Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,  
Chen,J., Chen,Z. and Han,Z.

Homo sapiens cDNA cda clones

Unpublished (2000)

Contact: Zequang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..482

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="cdABEB10"

/tissue\_type="pheochromocytoma"

/dev\_stage="Adult"

/lab\_host="BM25.8"

/clone\_lib="cda"

```

ORIGIN
/note="vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

Query Match      100.0%; Score 20; DB 9; Length 482;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
        |||||||
DB      180 CAGACGCATAGACCAACAGG 199

RESULT 24
BF819626/c

LOCUS      BF819626
DEFINITION BF81-RT0028-101100-002-b04 RT0028 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF819626
VERSION    BF819626.1 GI:12157598
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 484)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
PubMed    10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/getacml2.p?tl=MR1&t2=MR1-RT0028-101100-002-b04&t3=2000-11-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 484.
FEATURES
            Location/Qualifiers
                1..484
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="RT0028"
                    /note="Organ: kidney tumor; Vector: puc18; Site1: SmaI;
                    Site2: SmaI; A mini-library was made by cloning products
                    derived from ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 - Ludwig Institute for Cancer Research)
                    profiles into the pUC 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."

ORIGIN

Query Match      100.0%; Score 20; DB 10; Length 484;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGACGCATAGACCAACAGG 20
        |||||||
DB      426 CAGACGCATAGACCAACAGG 407

```



```

LOCUS       CD698819               514 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION   EST15342 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD698819
VERSION     CD698819.1 GI:32227504
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 514)
AUTHORS     Liu,X.-O., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.

FEATURES             Location/Qualifiers
     source           1..514
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /notes="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 514;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 247 CAGACGCATAGACCAACAGG 266

RESULT 27
LOCUS       BM769742               515 bp      mRNA      linear      EST 04-MAR-2002
DEFINITION   K-EST0053050 S14K402 Homo sapiens cDNA clone S14K402-25-B02 5',
            mRNA sequence.
ACCESSION   BM769742
VERSION     BM769742.1 GI:19099357
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 515)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boseun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 25 row: B column: 02
            High quality sequence stop: 515.
            Location/Qualifiers
                 1..515
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"

```

```

/clone="S14K402-25-B02"
/cell_line="K402"
/lab_host="T0p10F"
/clone_lib="S14K402"
/notes="Organ: Stomach; Vector: pTZ19RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match      100.0%; Score 20; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 190 CAGACGCATAGACCAACAGG 209

RESULT 28
LOCUS       BG541135               519 bp      mRNA      linear      EST 03-APR-2001
DEFINITION   602569911F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694379 5',
            mRNA sequence.
ACCESSION   BG541135
VERSION     BG541135.1 GI:13533368
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 519)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-x@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1518 row: e column: 04
            High quality sequence stop: 513.
            Location/Qualifiers
                 1..519
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:4694379"
                 /lab_host="NIH_MGC_77"
                 /clone_lib="NIH_MGC_77"
                 /notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                 SfiI (ggccattcgcc); Site_2: SfiI (ggccattcgcc); 5' and
                 3' adaptors were used in cloning as follows: 5' adaptor
                 sequence: 5'-CAGGCCCATATGGCC-3' and 3' adaptor sequence:
                 5'-ATTCTAGAGCCGAGGCGGCACATG-dT(30)BN-3' (where B = A,
                 C, or G and N = A, C, G, or T). Average insert size 1.9
                 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts

```

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 519;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAGCATAGACCAACAGG 20  
|||||  
Db 213 CAGAGCATAGACCAACAGG 232

## RESULT 29

CD705087 526 bp mRNA linear EST 25-JUN-2003  
LOCUS EST21614 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
DEFINITION CD705087  
ACCESSION CD705087.1 GI:32235717  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 526)  
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx

JOURNAL Unpublished (2003)  
COMMENT Contact: YiXin Zeng  
Cancer Center

Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@zsums.edu.cn.

## FEATURES

source  
1..526  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 526;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAGCATAGACCAACAGG 20  
|||||  
Db 251 CAGAGCATAGACCAACAGG 280

## RESULT 30

CD695435 540 bp mRNA linear EST 25-JUN-2003  
LOCUS EST11958 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
DEFINITION CD695435  
ACCESSION CD695435.1 GI:32220983  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 540)  
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: YiXin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@zsums.edu.cn.

## FEATURES

source  
1..540  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 540;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAGCATAGACCAACAGG 20  
|||||  
Db 220 CAGAGCATAGACCAACAGG 239

## RESULT 31

BE874055 544 bp mRNA linear EST 20-OCT-2000  
LOCUS 601484433F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3886761 5', mRNA sequence.  
DEFINITION BE874055  
ACCESSION BE874055  
VERSION BE874055.1 GI:10322831  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 544)  
AUTHORS NIH-MGC Http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: DCTD/BTP/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9664 row: b column: 10  
High quality sequence stop: 542.

## FEATURES

source  
1..544  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3886761"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_69"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; Nct1; Site2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.1 Kb. Library constructed by Life Technologies."

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 544;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

VERSION      AW406086.1  GI:6925107
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        (bases 1 to 547)
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Eco RI site shown at the beginning of the sequence.
              Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
              cDNA Library Preparation: M.B. Soares Lab
              cDNA Library Arrayed by: M.B. Soares Lab
              DNA Sequencing by: M.B. Soares Lab
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Seq primer: M13 Forward.
FEATURES     Location/Qualifiers
              1..547
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:3060714"
               /tissue_type="lymph"
               /cell_type="germinal center B cells"
               /cell_line="MGC85"
               /lab_host="DH10B (LTI)"
               /clone_lib="NIH_MGC_37"
               /notes="Vector: pT73-Pac; Site 1: NotI; Site 2: Eco RI;
               Constructed from size fractionated cytoplasmic mRNA
               (1.5-2.5kb). Directionally cloned. Cells provided by Louis
               M. Staudt, Ph.D. Library preparation by Maria de Fatima
               Bonalco, Ph.D. and M. Bento Soares, Ph.D."
ORIGIN
Query Match      100.0%; Score 20; DB 10; Length 547;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 176 CAGACGCATAGACCAACAGG 195
    |||||

RESULT 35
CD706950
LOCUS          547 bp mRNA linear EST 25-JUN-2003
DEFINITION    EST23477 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION     CD706950
VERSION       CD706950.1 GI:32237580
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     Liu X.-O., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
              Zeng, Y.-X.
AUTHORS       Transcriptional Gene Expression Profile of Human Nasopharynx
              Unpublished (2003)
              Contact: Yixin Zeng
              Cancer Center
              Sun Yat-sen University
              651 Dongfeng Road East, Guangzhou 510060, China
              Tel: 86-1360-9770-743
              Fax: 86-20-8775-4506
              Email: yxzeng@gzsums.edu.cn.
              Location/Qualifiers
              1..547

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 547;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 218 CAGACGCATAGACCAACAGG 237
    |||||

RESULT 36
BG535978
LOCUS          557 bp mRNA linear EST 03-APR-2001
DEFINITION    602564150F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4688604 5',
              mRNA sequence.
ACCESSION     BG535978
VERSION       BG535978.1 GI:13527523
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS       National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: CLONTECH Laboratories, Inc.
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM1503 row: d column: 13
              High quality sequence stop: 557.
              Location/Qualifiers
              1..557
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:4688604"
               /lab_host="DH10B (T1 phage-resistant)"
               /clone_lib="NIH_MGC_77"
               /note="Organ: lung; Vector: pDNR-LTB (Clontech); Site 1:
               SfiI (ggccgctggcc); Site 2: SfiI (ggccattggcc); 5' and
               3' adaptors were used in cloning as follows: 5' adaptor
               sequence: 5'-CAGCGCATATGCG-3' and 3' adaptor sequence:
               5'-ATTCTAGAGCGCGAGCGCGCATG-DT(30)EN-3' (where B = A,
               C, or G and N = A, C, G, or T). Average insert size 1.9
               kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
               by PCR. This library was enriched for full-length clones
               and was constructed by Clontech Laboratories (Palo Alto,
               CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 20; DB 12; Length 557;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCATAGACCAACAGG 20
    |||||
Db 236 CAGACGCATAGACCAACAGG 255
    |||||

```

```

RESULT 37
CD687141
LOCUS      CD687141          557 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION EST3662 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD687141
VERSION     CD687141.1  GI:32204715
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 Dongfeng Road East, Guangzhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES    source
            1..557
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="normal nasopharynx"
            /clone_lib="human nasopharynx"
            /notes="ESTs generated from a normal nasopharynx cDNA
            library from southern Chinese"
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CAGACGCATAGACCAACAGG 20
    |||
Db   140 CAGACGCATAGACCAACAGG 159
    |||

RESULT 38
CD700948
LOCUS      CD700948          559 bp      mRNA      linear      EST 25-JUN-2003
DEFINITION EST17504 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD700948
VERSION     CD700948.1  GI:32231610
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 Dongfeng Road East, Guangzhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES    source
            1..559
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="normal nasopharynx"

```

```

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 559;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CAGACGCATAGACCAACAGG 20
    |||
Db   254 CAGACGCATAGACCAACAGG 273
    |||

RESULT 39
CB265399
LOCUS      CB265399          563 bp      mRNA      linear      EST 20-FEB-2003
DEFINITION 1004304 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
            cDNA 5', mRNA sequence.
ACCESSION  CB265399
VERSION     CB265399.1  GI:28439987
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
AUTHORS     Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE       EST analysis of human adipose gene expression
JOURNAL     Unpublished (2002)
COMMENT     Contact: Gong Da-Wei
            Division of Endocrinology, Diabetes and Nutrition
            University of Maryland
            660 Redwood St, RM497, Baltimore, MD 21201, USA
            Tel: 410 706 1672
            Fax: 410 706 1622
            Email: dgong@medicine.umaryland.edu
            PCR Primers
            FORWARD: CTCGGGACGCGCCATTTGTGTGGT
            BACKWARD: AATACGACTCATATAGGCGCATTTGG
            Seq primer: GTTGGTACCCGGGAAATTC.
FEATURES    Location/Qualifiers
            1..563
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /sex="Male and Female"
            /tissue_type="Adipose"
            /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
            /note="Vector: lambdaTriplex"
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 563;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  CAGACGCATAGACCAACAGG 20
    |||
Db   194 CAGACGCATAGACCAACAGG 213
    |||

RESULT 40
CA942442
LOCUS      CA942442          571 bp      mRNA      linear      EST 30-DEC-2002
DEFINITION ir58a06.y1 HR8s islet Homo sapiens cDNA clone IMAGE:6549227 5',
            similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY
            ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.
ACCESSION  CA942442
VERSION     CA942442.1  GI:27430922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
AUTHORS  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 571)  
Melton, D., Brown, J., Kenry, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Killer, L., Ma, M., Pape, D., Wyllie, T., Martin, J., Blisstein, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R.,  
Williams, T., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other\_RSTS: ir58a06.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -400P from Gibco  
High quality sequence stop: 487.

FEATURES  
source

1. .571  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6549227"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site: 1:  
NotI; Site: 2: XhoI; cDNA made by oligo-dT priming.  
Size: selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 571;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 194 CAGACGCATAGACCAACAGG 213

RESULT 41  
BM831052  
LOCUS K-EST0104933 S14K402s1 Homo sapiens cDNA clone S14K402s1-18-C12 5',  
EST 06-MAR-2002  
DEFINITION mRNA sequence.  
BM831052  
VERSION BM831052.1 GI:19187461  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 18 row: C column: 12  
High quality sequence stop: 579.  
Location/Qualifiers  
1. .579

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S14K402s1-18-C12"  
/cell\_line="K402"  
/lab\_host="Top10P"  
/clone\_lib="S14K402s1"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;  
Site: 2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including BcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10P by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library. After analyzing and  
sequencing about 2,000 ~ 3,000 colonies in original cDNA  
library, the abundant cDNAs were selected and amplified by  
PCR reaction using vector region primer including T7  
promoter as 5' primer and N(dT)14 as 3' primer. The PCR  
products were used as template for synthesis of  
biotinylated single stranded RNA by in vitro transcription  
reaction. The synthesized RNA probes were hybridized with  
antisense single stranded cDNAs prepared from original  
library and incubated with avidin-gel. After removing  
DNA-RNA hybrids by centrifuge, the subtracted cDNA  
libraries were constructed by transformation of the  
remaining DNA into competent cells E. coli Top10P with  
electroporation method."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 579;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 150 CAGACGCATAGACCAACAGG 169

RESULT 42  
BM737984  
LOCUS K-EST0002067 S1SNU5 Homo sapiens cDNA clone S1SNU5-25-B01 5', mRNA  
EST 01-MAR-2002  
DEFINITION sequence.  
BM737984 582 bp  
VERSION BM737984.1 GI:19059313  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.

# TITLE JOURNAL COMMENT

21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 25 row: B column: 01  
High quality sequence stop: 582.

## FEATURES

source

1. 582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SISNU5-25-B01"  
/sex="p"  
/tissue\_type="Ascites"  
/cell\_type="lymphoblast-like"  
/cell\_line="SNU-5"  
/lab\_host="Top10"  
/clone\_lib="SISNU5"  
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transfection of  
competent cells E. coli Top10<sup>+</sup> by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 582;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 213 CAGACGCATAGACCAACAGG 232

## RESULT 43

AV706521  
LOCUS AV706521 ADB Homo sapiens cDNA clone ADBAve11 5', mRNA linear EST 09-OCT-2000  
DEFINITION AV706521  
ACCESSION AV706521  
VERSION AV706521.1 GI:10723800

KEYWORDS EST

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 585)

REFERENCE Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,

Xiao, H., Xu, X., Li, N., Qian, B., Liu, P., Qu, J., Gao, X., Cheng, Z.,

Xu, Z., Zeng, B., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,

Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.

Homo sapiens cDNA ADB clones

Unpublished (2000)

Contact: Zequang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex. 45)

# FEATURES source

1. 585  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ADBAVE11"  
/tissue\_type="Adrenal gland"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="ADB"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

Db 264 CAGACGCATAGACCAACAGG 283

## RESULT 44

BU783392  
LOCUS BU783392 585 bp mRNA linear EST 11-OCT-2002  
DEFINITION in03f10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123426  
5', similar to SW:HA2Q HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY  
ANTIGEN, DP ALPHA CHAIN PRECURSOR, mRNA sequence.

ACCESSION BU783392

VERSION BU783392.1 GI:23827516

KEYWORDS EST

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 585)

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Renko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McGann, R., Cole, R., Tsagarisvili, R.,

Williams, J., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Other ESTs: in03f10.x1

Unpublished (2000)

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to

pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 429.

Location/Qualifiers

1. 585

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6123426"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:

XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 585;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 147 CAGACGCATAGACCAACAGG 166

RESULT 45  
BI911442 588 bp mRNA linear EST 16-OCT-2001  
LOCUS 603063356f1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212669 5',  
DEFINITION mRNA sequence.

ACCESSION BI911442  
VERSION BI911442.1 GI:16175182  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 588)

AUTHORS NIH-MGC http://rgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LIA01533 row: p column: 14  
High quality sequence start: 18  
High quality sequence stop: 585.

## FEATURES

Location/Qualifiers  
1..588  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5212669"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC 118"

/note="vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 588;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20  
|||||  
Db 205 CAGACGCATAGACCAACAGG 224

## RESULT 46

BM876262

## LOCUS

## DEFINITION

BM876262 592 bp mRNA linear EST 07-MAR-2002  
ij60c02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5635202  
3' similar to SW:HA2Q.HUMAN P20036 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP ALPHA CHAIN PRECURSOR ;, mRNA sequence.

## ACCESSION

BM876262

## VERSION

BM876262.1 GI:19243928

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 592)

## AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brastelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marx,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,B., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.

## TITLE

Endocrine Pancreas Consortium

## JOURNAL

Unpublished (2000)

## COMMENT

Other ESTs: ij60c02.y1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@wustl.edu)  
Possible reversed clone: similarity on wrong strand  
Seq primer: -400P from Gibco  
High quality sequence stop: 482.

## FEATURES

Location/Qualifiers

1..592  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5635202"  
/tissue\_type="insulinoma"  
/lab\_host="Human insulinoma"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).

Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 592;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGACGCATAGACCAACAGG 20

|||||

Db 188 CAGACGCATAGACCAACAGG 207



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 634.447 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819B-40

Perfect score: 21  
Sequence: 1 cccgtgtggctatcggtctg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 180 summaries

Database :

GenEmbl: \*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_cm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_man:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htg\_hum:\*

40: em\_htg\_mus:\*

41: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	21	100.0	189	9	MMDPALP	Z32411 M.mulatta (
C 2	21	100.0	192	9	HUNMHCIITF	L31624 Human MHC c
C 3	21	100.0	246	9	AF346471	AF346471 Homo sapi
C 4	21	100.0	246	9	HS0103X2	X83390 H.sapiens M
C 5	21	100.0	246	9	HS0103X2	X78198 H.sapiens H
C 6	21	100.0	249	9	HS02PALRK	X96984 Homo sapien
C 7	21	100.0	267	6	AX237167	AX237167 Sequence
C 8	21	100.0	267	6	AX237352	AX237352 Sequence
C 9	21	100.0	268	9	HS01ADPAX	AX237066 Sequence
C 10	21	100.0	272	6	AX237066	AX237066 Sequence
C 11	21	100.0	279	9	AF076284	AF076284 Homo sapi
C 12	21	100.0	279	9	AF076285	AF076285 Homo sapi
C 13	21	100.0	286	9	HS087556	U87556 Homo sapien
C 14	21	100.0	287	9	AF015295	AF015295 Homo sapi
C 15	21	100.0	294	6	AX237304	AX237304 Sequence
C 16	21	100.0	294	6	AX237554	AX237554 Sequence
C 17	21	100.0	466	6	BD058339	BD058339 Secreter
C 18	21	100.0	476	6	AX884252	AX884252 Sequence
C 19	21	100.0	476	6	BD023862	BD023862 Sequence
C 20	21	100.0	576	6	AX884251	AX884251 Sequence
C 21	21	100.0	576	6	BD023861	BD023861 Sequence
C 22	21	100.0	661	6	E00485	E00485 DNA sequenc
C 23	21	100.0	661	6	I03086	I03086 Sequence 5
C 24	21	100.0	690	9	S40633	S40633 HLA class I
C 25	21	100.0	818	9	HUNMHDPDA	M27487 Homo sapien
C 26	21	100.0	1048	9	HSSBAL	X00457 Human mRNA
C 27	21	100.0	1140	6	I03088	I03088 Sequence 7
C 28	21	100.0	1201	6	E00484	E00484 DNA sequenc
C 29	21	100.0	1201	6	I03006	I03006 Sequence 6
C 30	21	100.0	1201	6	I03423	I03423 Sequence 6
C 31	21	100.0	1259	6	AX552229	AX552229 Sequence
C 32	21	100.0	12986	9	HUNMHDRCC02	M23904 Human MHC c
C 33	21	100.0	14646	9	HS01ASBA	X03100 Human HLA-S
C 34	21	100.0	64380	9	AL805913	AL805913 Human DNA
C 35	21	100.0	106728	9	AL805913	AL805913 Human DNA
C 36	21	100.0	124989	2	AC011866	AC011866 Homo sapi
C 37	21	100.0	181228	2	AC011866	AC011866 Homo sapi
C 38	21	100.0	187964	9	AL662824	AL662824 Human DNA
C 39	20	95.2	222	9	AF074848	AF074848 Homo sapi
C 40	20	95.2	244	9	AF098794	AF098794 Homo sapi
C 41	20	95.2	1480	6	AX780118	AX780118 Sequence
C 42	20	95.2	1480	6	AX780119	AX780119 Sequence
C 43	19.4	92.4	189	9	AF026693	AF026693 Pan trogl
C 44	19.4	92.4	189	9	AF026694	AF026694 Pan trogl
C 45	19.4	92.4	189	9	AF026696	AF026696 Pongo pyg
C 46	19.4	92.4	189	9	AF026700	AF026700 Saimiri s
C 47	19.4	92.4	189	9	AF026703	AF026703 Macaca fa
C 48	19.4	92.4	189	9	AF026704	AF026704 Macaca fa
C 49	19.4	92.4	189	9	AF026705	AF026705 Macaca mu
C 50	19.4	92.4	189	9	AF026706	AF026706 Papio ham
C 51	19.4	92.4	189	9	AF026707	AF026707 Pan trogl
C 52	19.4	92.4	189	9	AF529200	AF529200 Aotus nan
C 53	19.4	92.4	189	9	HUNMHAMALX	M83908 Human MHC c
C 54	19.4	92.4	226	9	HS02PALAI	X79477 H.sapiens M
C 55	19.4	92.4	246	9	HS0201X2	X82394 H.sapiens M
C 56	19.4	92.4	252	9	HS248473	Z48473 H.sapiens H
C 57	19.4	92.4	257	9	AF165160	AF165160 Homo sapi
C 58	19.4	92.4	258	9	AF118120	AF118120 Homo sapi
C 59	19.4	92.4	265	9	AF092049	AF092049 Homo sapi
C 60	19.4	92.4	326	9	AF013767	AF013767 Homo sapi
C 61	19.4	92.4	396	6	BD058253	BD058253 Secreter
C 62	19.4	92.4	1201	9	BC009956	BC009956 Homo sapi
C 63	19	90.5	3888	10	AY082609	AY082609 Rattus no
C 64	19	90.5	4233	6	AR123273	AR123273 Sequence
C 65	19	90.5	4254	6	AX401748	AX401748 Sequence

66	19	90.5	4254	6	AX827531	AX827531 Sequence	C 139	17.4	82.9	105604	9	AC084013	Homo sapi
67	19	90.5	4254	10	RATMDRM	M81855 Rat mdr mkn	140	17.4	82.9	106642	9	AL662781	Human DNA
68	19	90.5	168634	2	AC133679	Rattus no	C 141	17.4	82.9	109774	2	AC015923	Homo sapi
69	19	90.5	236777	2	AC094362	Rattus no	C 142	17.4	82.9	110000	2	AC106723	Continuation (3 of
70	18.4	87.6	189	9	AF026695	Pongo pyg	C 143	17.4	82.9	113799	9	AC012502	Homo sapi
71	18.4	87.6	189	9	AF026697	Pongo pyg	C 144	17.4	82.9	120112	9	AC012355	Homo sapi
72	18.4	87.6	189	9	AF026701	Gorilla g	C 145	17.4	82.9	121158	9	AC092977	Homo sapi
73	18.4	87.6	189	9	AF026702	Gorilla g	146	17.4	82.9	122352	9	HS104401	Human DNA
74	18.4	87.6	192	9	HUMHCHIAZ	L11643 Human MHC c	C 147	17.4	82.9	127703	9	AL357034	Human DNA
75	18.4	87.6	198	9	HUMHCHIB	M83909 Human MHC c	C 148	17.4	82.9	129740	9	AC091525	Homo sapi
76	18.4	87.6	198	9	HUMHCHIBEX	M83906 Human MHC c	C 149	17.4	82.9	130110	9	AL731777	Human DNA
77	18.4	87.6	198	9	HUMHCHIT4	M83907 Human MHC c	C 150	17.4	82.9	130351	2	AC024911	Homo sapi
78	18.4	87.6	214	9	HUMHCHIT4X	L11641 Human MHC c	C 151	17.4	82.9	132229	9	HS447N6	Human DNA
79	18.4	87.6	220	9	HUMHCHIT4Y	L11642 Human MHC c	C 152	17.4	82.9	135484	2	AL356301	Homo sapi
80	18.4	87.6	226	9	HSDFALIB	X79478 H.sapiens M	C 153	17.4	82.9	135978	2	CNS01DV8	Homo sapi
81	18.4	87.6	226	9	HSDFALIB	X79475 H.sapiens M	C 154	17.4	82.9	137590	2	AC027567	Homo sapi
82	18.4	87.6	3161	9	HSDFALIB	X79476 H.sapiens M	C 155	17.4	82.9	138235	2	AL157393	Human DNA
83	18.4	87.6	36171	9	HSDFALIB	AK126094 Homo sapi	C 156	17.4	82.9	141854	2	AC021595	Homo sapi
84	18.4	87.6	39756	9	HSDFALIB	AK126094 Homo sapi	C 157	17.4	82.9	142008	9	AC025257	Homo sapi
85	18.4	87.6	39756	9	HSDFALIB	Z54246 Human DNA s	C 158	17.4	82.9	142439	9	AC008749	Homo sapi
86	18.4	87.6	141287	2	AC093404	AC093404 Pan trogl	C 159	17.4	82.9	143035	9	AC010082	Homo sapi
87	18.4	87.6	145151	9	AL645949	AL645949 Human DNA	C 160	17.4	82.9	143244	2	AC087655	Homo sapi
88	18.4	87.6	164590	9	AL390059	AL390059 Human DNA	C 161	17.4	82.9	143417	9	AC025157	Homo sapi
89	18.4	87.6	200348	9	AC005516	AC005516 Homo sapi	C 162	17.4	82.9	143552	9	AC105021	Homo sapi
90	18.4	87.6	235097	2	AC110877	Mus muscu	C 163	17.4	82.9	145957	9	AC024566	Homo sapi
91	17.8	84.8	189	9	AF026698	Saimiri s	C 164	17.4	82.9	147109	9	AL358779	Human DNA
92	17.8	84.8	189	9	AF026699	Saimiri s	C 165	17.4	82.9	147156	2	AC025742	Homo sapi
93	17.8	84.8	222	9	AF074847	AF074847 Homo sapi	C 166	17.4	82.9	147359	9	AP002345	Homo sapi
94	17.8	84.8	222	9	AF074847	AB084158 Cavia por	C 167	17.4	82.9	148005	2	AC024944	Homo sapi
95	17.8	84.8	3056	10	AF027333	AF027333 Rattus no	C 168	17.4	82.9	148193	9	AC027347	Homo sapi
96	17.8	84.8	3332	8	AK110088	AK110088 Rytza sat	C 169	17.4	82.9	148295	9	AC090510	Homo sapi
97	17.8	84.8	3511	10	RATTTT1	D38035 Rattus norv	C 170	17.4	82.9	148476	9	AC130313	Homo sapi
98	17.8	84.8	11172	1	MMU19755	U19755 Mus domesti	C 171	17.4	82.9	149102	9	AC108095	Homo sapi
99	17.8	84.8	17976	1	AE006939	AE006939 Mycobacte	C 172	17.4	82.9	149416	2	AC021985	Homo sapi
100	17.8	84.8	78132	2	AC021703	AC021703 Homo sapi	C 173	17.4	82.9	149731	9	HS162013	Human DNA
101	17.8	84.8	149866	2	AC069256	AC069256 Homo sapi	C 174	17.4	82.9	150339	9	AC099515	Homo sapi
102	17.8	84.8	153251	2	AC117653	AC117653 Mus muscu	C 175	17.4	82.9	150401	9	AC091907	Homo sapi
103	17.8	84.8	154251	2	AC092446	AC092446 Homo sapi	C 176	17.4	82.9	151348	9	AC015807	Homo sapi
104	17.8	84.8	168991	2	CNS01DUA	AL133233 Human chr	C 177	17.4	82.9	152449	2	AC080175	Homo sapi
105	17.8	84.8	171415	9	AC093807	AC093807 Homo sapi	C 178	17.4	82.9	153089	9	AC146510	Pan trogl
106	17.8	84.8	176021	2	AC026582	AC026582 Homo sapi	C 179	17.4	82.9	153496	2	AC087044	Homo sapi
107	17.8	84.8	184688	9	AC092944	AC092944 Homo sapi	180	17.4	82.9	153938	2	AC080050	Homo sapi
108	17.8	84.8	190000	2	AC006391	AC006391 Homo sapi							
109	17.8	84.8	194277	9	AC096865	AC096865 Homo sapi							
110	17.8	84.8	204098	2	EX088705	EX088705 Danio rer							
111	17.8	84.8	213462	2	AC119365	AC119365 Rattus no							
112	17.8	84.8	215514	2	AC040982	AC040982 Mus muscu							
113	17.8	84.8	219313	2	AC120967	AC120967 Rattus no							
114	17.8	84.8	228931	2	AC131355	AC131355 Rattus no							
115	17.8	84.8	248124	2	AC096607	AC096607 Rattus no							
116	17.8	84.8	255172	2	AC120724	AC120724 Rattus no							
117	17.8	84.8	257339	2	AC128869	AC128869 Rattus no							
118	17.8	84.8	257619	2	AC114512	AC114512 Rattus no							
119	17.8	84.8	324050	1	EX248335	EX248335 Mycobacte							
120	17.8	84.8	342416	1	EX42573	EX42573 Mycobacte							
121	17.4	82.9	36353	9	HS9F14	AL096177 Homo sapi							
122	17.4	82.9	44262	9	AL353808	AL353808 Human DNA							
123	17.4	82.9	52664	9	AL732432	AL732432 Human DNA							
124	17.4	82.9	62104	2	AC024919	AC024919 Homo sapi							
125	17.4	82.9	64744	2	AC090694	AC090694 Homo sapi							
126	17.4	82.9	69471	2	AC117400	Continuation (4 of							
127	17.4	82.9	72978	9	AL359759	AL359759 Human DNA							
128	17.4	82.9	74874	9	AC110056	AC110056 Homo sapi							
129	17.4	82.9	80611	9	HS323824	AL022157 Human DNA							
130	17.4	82.9	81091	9	AL645495	AL645495 Human DNA							
131	17.4	82.9	86563	2	AC139663	AC139663 Homo sapi							
132	17.4	82.9	88614	2	AC026865	AC026865 Homo sapi							
133	17.4	82.9	90171	9	AC095040	AC095040 Homo sapi							
134	17.4	82.9	92346	9	AC074289	AC074289 Homo sapi							
135	17.4	82.9	96209	9	AX895455	AX895455 Sequence							
136	17.4	82.9	96596	6	AX695456	AL096770 Human DNA							
137	17.4	82.9	97392	9	HSAL150A6	AL953706 Human DNA							
138	17.4	82.9	104679	9	AL353706								

## ALIGNMENTS

RESULT 1  
MMDPALF/c  
LOCUS MMDPALF 189 bp DNA linear PRI 26-JUL-1995  
DEFINITION M. mulatta (clone Manu-DPA1\*0101) Mhc DP-alpha gene encoding major histocompatibility complex.

ACCESSION Z32411.1 GI:471260  
VERSION Z32411.1  
KEYWORDS major histocompatibility complex.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta

REFERENCE 1 (bases 1 to 189)  
AUTHORS Slierendregt, B.L., Otting, N., Kenter, M. and Bontrop, R.E.  
TITLE Allelic diversity at the Mhc-DP locus in rhesus macaques (Macaca mulatta)

JOURNAL Immunogenetics 41 (1), 29-37 (1995)  
MEDLINE 95104902  
PUBMED 7806271

REFERENCE 2 (bases 1 to 189)  
AUTHORS Slierendregt, B.B.  
TITLE Direct Submission

Submitted (07-APR-1994) Bastiaan B.L. Slierendregt, Immunobiology, Biomedical Primate, Research Center-TNO, Lange Kleiweg 151, Rijswijk, Zuid-Holland, 2288 GJ, The Netherlands

```

FEATURES
  Source
    Location/Qualifiers
      1..189
        /organism="Macaca mulatta"
        /mol_type="genomic DNA"
        /isolate="2849"
        /db_xref="taxon:9544"
        /clone="Mamu-DPA1*0101"
        /cell_type="Lymphocyte"
        /tissue_type="Blood"
      1..189
        /gene="Mhc DP-alpha"
        <1..>189
        /gene="Mhc DP-alpha"
        /note="exon 2"
        /codon_start=1
        /product="major histocompatibility complex"
        /protein_id="CAA83527.1"
        /db_xref="GI:471261"
        /db_xref="GOA:Q30617"
        /db_xref="SPTREMBL:Q30617"
        /translation="FVQHRPTGGEFDEMFYDLKKETVHLEBFQARSPF
AOGGLANIALNNLNLTIL"
      ORIGIN
        Query Match 100.0%; Score 21; DB 9; Length 189;
        Best Local Similarity 100.0%; Pred. No. 0.53; 0; Indels 0; Gaps 0;
        Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 CCTGTGTCATGCGCTG 21
      |||||
  Db 27 CCTGTGTCATGCGCTG 7

RESULT 2
HUMHCHITF/c
LOCUS HUMHCHITF 192 bp DNA linear PRI 03-MAY-1995
DEFINITION Human MHC class II gene (DPA1*TF).
ACCESSION L31624
VERSION L31624.1 GI:598150
KEYWORDS cell surface glycoprotein; class II gene; integral membrane
protein; major histocompatibility complex.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Meyer,C.G., May,J., Spaube,D. and Schnittger,L.
AUTHORS DPA1*02012: a DPA1*0201-related Mhc class II allele in west Africa
JOURNAL Immunogenetics 40 (4), 309 (1994)
MEDLINE 94364641
PUBMED 8082895
COMMENT Original source text: Homo sapiens male blood DNA.
FEATURES
  Source
    Location/Qualifiers
      1..192
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /sex="male"
        /cell_type="lymphocyte"
        /tissue_type="blood"
      ORIGIN
        Query Match 100.0%; Score 21; DB 9; Length 192;
        Best Local Similarity 100.0%; Pred. No. 0.53; 0; Indels 0; Gaps 0;
        Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 CCTGTGTCATGCGCTG 21
      |||||
  Db 33 CCTGTGTCATGCGCTG 13

RESULT 3
AF346471/c
LOCUS AF346471 246 bp DNA linear PRI 27-MAR-2001
DEFINITION Homo sapiens MHC class II antigen (HLA-DPA1) gene, HLA-DPA1-new
allele, partial cds.
ACCESSION AF346471
VERSION AF346471.1 GI:13448661
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Grams,S.E., Begovich,A. and Mangaccat,J.
AUTHORS 1 (bases 1 to 246)
TITLE One new DPA1 Allele
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 246)
AUTHORS Grams,S.E., Begovich,A. and Mangaccat,J.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2001) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
FEATURES
  Source
    Location/Qualifiers
      1..246
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="6"
        /map="6p"
        <1..>246
        /gene="HLA-DPA1"
        /alleles="new"
        <1..>246
        /gene="HLA-DPA1"
        /product="MHC class II antigen"
        <1..>246
        /gene="HLA-DPA1"
        /codon_start=3
        /product="MHC class II antigen"
        /protein_id="AAK27152.1"
        /db_xref="GI:13448662"
        /translation="DHVSTIAFVQTHRPTGEFDEMFYDLKKETVHLE
FGRAFSFAOGLANIALNNLNLTILQRSNHTQATN"
      1..246
        /gene="HLA-DPA1"
        /number=2
      ORIGIN
        Query Match 100.0%; Score 21; DB 9; Length 246;
        Best Local Similarity 100.0%; Pred. No. 0.55;
        Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 CCTGTGTCATGCGCTG 21
      |||||
  Db 53 CCTGTGTCATGCGCTG 33

RESULT 4
HS0103X2/c
LOCUS HS0103X2 246 bp DNA linear PRI 14-JUL-1995
DEFINITION H.sapiens MHC class II HLA-DPA1*0103 Gene (exon 2).
ACCESSION X82390
VERSION X82390.1 GI:565028
KEYWORDS MHC class II HLA DPA1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 246)
TITLE Roemuller,E.H., Bouwens,A.G., van Oort,E., Versluis,I.P.,
Marsh,S.G., Bodmer,J.G. and Tilius,M.G.
Sequencing-based typing reveals new insight in HLA-DPA1
polymorphism
JOURNAL Tissue Antigens 45 (1), 57-62 (1995)
MEDLINE 95242313
PUBMED 7725312

```

```

REFERENCE      2 (bases 1 to 246)
AUTHORS        Rozenmuller,E.H.
TITLE          Direct Submission
JOURNAL        Submitted (19-OCT-1994) E.H. Rozenmuller, Diagnostic DNA Laboratory,
               University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
               Utrecht, NETHERLANDS
COMMENT        Related sequences: S52453 and D14344.
               Related sequences: S52453 and D14344.
FEATURES       Location/Qualifiers
               1..246
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="6"
               /haplotype="DPA1*0103"
               /cell_line="LB and LG2"
gene           1..246
               /gene="MHC class II HLA-DPA1 gene"
exon           1..246
               /gene="MHC class II HLA-DPA1 gene"
               /number=2

ORIGIN
Query Match      100.0%; Score 21; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATCGGCTG 21
        |||||
Db       53 CCTGTTGGTCTATCGGCTG 33

RESULT 5
HSHLADPA1/c
LOCUS          HSHLADPA1      246 bp      DNA      linear      PRI 14-JUL-1995
DEFINITION     Homo sapiens HLA-DPA1 gene, exon 2.
ACCESSION      X781198
VERSION        X781198.1 GI:461355
KEYWORDS       DPA1*01new; HLA-DPA1 gene.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 246)
AUTHORS        Rozenmuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,
               Marsh,S.G., Boemer,J.G. and Tilius,M.G.
TITLE          Sequencing-based typing reveals new insight in HLA-DPA1
               Polymorphism
JOURNAL        Tissue Antigens 45 (1), 57-62 (1995)
MEDLINE        95242313
PubMed        7725312
REFERENCE      2 (bases 1 to 246)
AUTHORS        Rozenmuller,E.H.
TITLE          Direct Submission
JOURNAL        Submitted (15-MAR-1994) E.H. Rozenmuller, Diagnostic DNA Lab,
               University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
               Utrecht, NETHERLANDS
FEATURES       Location/Qualifiers
               1..246
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /isolate="SK"
               /db_xref="taxon:9606"
               /chromosome="6"
               /haplotype="DPA1*01"
gene           1..246
               /gene="HLA-DPA1"
exon           1..246
               /gene="HLA-DPA1"
               /product="DPA1*01new"
               /number=2

ORIGIN

```

```

Query Match      100.0%; Score 21; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATCGGCTG 21
        |||||
Db       53 CCTGTTGGTCTATCGGCTG 33

RESULT 6
HSDPAIRK/c
LOCUS          HSDPAIRK      249 bp      DNA      linear      PRI 24-NOV-2000
DEFINITION     Homo sapiens HLA-DPA1 gene, exon 2, isolate DNA RK.
ACCESSION      X96984
VERSION        X96984.1 GI:1590762
KEYWORDS       antigen; MHC; MHC class II; MHC class II DPA1 gene.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS        May,J., Krestchmer,C., Schnittger,L., Striecker,R., Kremener,P.G.
               and Meyer,C.G.
TITLE          DPA1*0105, a novel DPA1 variant in a negrois population
JOURNAL        Tissue Antigens 48, 693-694 (1996)
REFERENCE      2 (bases 1 to 249)
AUTHORS        Meyer,C.G.
TITLE          Direct Submission
JOURNAL        Submitted (29-MAR-1996) C.G. Meyer, Insitute of Tropical Medicine
               Berlin, Mol. Biol., Engeldamm 62, 10179 Berlin, FRG
               Location/Qualifiers
               1..249
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /isolate="DNA RK from Gabon"
               /db_xref="taxon:9606"
               /chromosome="6"
               /tissue type="blood"
               /dev stage="adult"
               10..784
               /product="MHC-class II DPA1 antigen"
               /number=2

exon

ORIGIN
Query Match      100.0%; Score 21; DB 9; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATCGGCTG 21
        |||||
Db       33 CCTGTTGGTCTATCGGCTG 13

RESULT 7
AX237167/c
LOCUS          AX237167      267 bp      DNA      linear      PAT 26-SEP-2001
DEFINITION     Sequence 143 from Patent WO0164886.
ACCESSION      AX237167
VERSION        AX237167.1 GI:15796721
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS        Gaiger,A., Algate,P.A. and Mannion,J.
TITLE          Compositions and methods for the detection, diagnosis and therapy
               of hematological malignancies
JOURNAL        Patent: WO 0164886-A 143 07-SEP-2001;
               CORIXA CORPORATION (US)
               Location/Qualifiers
               1..267
               /organism="Homo sapiens"

```

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 181 CCTGTGGTCTATCGGCTG 161

RESULT 8  
AX237352/c  
LOCUS AX237352 267 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 328 from Patent WO0164886.  
ACCESSION AX237352  
VERSION AX237352.1 GI:15796906

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.  
TITLE Compositions and methods for the detection, diagnosis and therapy  
of hematological malignancies  
JOURNAL Patent: WO 0164886-A 328 07-SEP-2001;  
CORIXA CORPORATION (US)

FEATURES  
source  
1..267  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 181 CCTGTGGTCTATCGGCTG 161

RESULT 9  
HSHLADPAX/c  
LOCUS HSHLADPAX 268 bp DNA linear PRI 14-SEP-1995  
DEFINITION H.sapiens HLA-DPA1 gene.  
ACCESSION X83610  
VERSION X83610.1 GI:987073

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Versluis, L.F., Verdunyn, W., Van der Zwan, A., Oudshoorn, M. and  
Tilanus, M.G.J.  
TITLE An update of the exon 2 sequence of the HLA-DPA1\*02012 allele

JOURNAL Tissue Antigens 46 (3 Pt 1), 206-207 (1995)  
MEDLINE 96097411  
PUBMED 8525481

REFERENCE  
2 (bases 1 to 268)  
Tilanus, M.G.J.

AUTHORS Direct Submission  
TITLE Submitted (20-DEC-1994) M.G.J. Tilanus, Diagnostic DNA Laboratory,  
Academic Hospital Utrecht, Heidelberglaan 100, PO Box 85500, 3508  
GA Utrecht, NETHERLANDS  
COMMENT Related sequence: I31624.

FEATURES  
Location/Qualifiers  
1..268

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/haplotype="DPA1\*02012"

intron  
gene  
19..264  
/gene="HLA-DPA1"  
exon  
19..264  
/gene="HLA-DPA1"  
intron  
265..268

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 71 CCTGTGGTCTATCGGCTG 51

RESULT 10  
AX237066/c  
LOCUS AX237066 272 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 42 from Patent WO0164886.  
ACCESSION AX237066  
VERSION AX237066.1 GI:15796620

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.  
TITLE Compositions and methods for the detection, diagnosis and therapy  
of hematological malignancies  
JOURNAL Patent: WO 0164886-A 42 07-SEP-2001;  
CORIXA CORPORATION (US)

FEATURES  
source  
1..272  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 186 CCTGTGGTCTATCGGCTG 166

RESULT 11  
AF076284/c  
LOCUS AF076284 279 bp DNA linear PRI 07-JUL-1999

DEFINITION Homo sapiens isolate 913 MHC class II antigen (HLA-DPA1) gene,  
partial cds.  
ACCESSION AF076284  
VERSION AF076284.1 GI:5381292

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Varney, M.D., Gavrilidis, A. and Abbott, W.  
TITLE DPAP Polymorphism in Polynesians  
JOURNAL Unpublished

REFERENCE  
2 (bases 1 to 279)  
Varney, M.D., Gavrilidis, A. and Abbott, W.

```

TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) Tissue Typing, Royal Melbourne Hospital,
Grattan Street, Parkville, Vic 3050, Australia
FEATURES
source
1..279
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="g13"
/db_xref="taxon:9606"
/chromosome="6"
<1..>279
/gene="HLA-DPA1"
<1..>279
/gene="HLA-DPA1"
<1..>279
/product="MHC class II antigen"
<1..>279
/gene="HLA-DPA1"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAD42927.1"
/db_xref="GI:5381293"
/translation="HVSYYAAFVOTHRPTGFMFPEFDEQFYVDLKKETVWHLERF
GQTFSEAQGLANIALINNNLTLIQRSNHTQAANGTPLYCLFLCSPT"
exon
1..279
/gene="HLA-DPA1"
/number=2

ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTTGGTCTATGCGTCTG 21
|||||
DB 48 CCTGTTGGTCTATGCGTCTG 28

RESULT 13
LOCUS HSU87556/c
DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2,
partial cds.
ACCESSION U87556
VERSION U87556.1 GI:2760313
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 286)
Steiner, L., Begovich, A. and Suraj, V.
Direct Submission
JOURNAL Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
COMMENT On Jan 8, 1998 this sequence version replaced gi:1842112.
FEATURES
source
1..286
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
<1..>286
/gene="HLA-D"
<16..>261
/gene="HLA-D"
/note="MHC class II HLA-DPA1 antigen"
/codon_start=3
/protein_id="AAB97110.1"
/db_xref="GI:2795772"
/translation="DHVSTYAAFVOTHRPTGFMFPEFDEQFYVDLKKETVWHLERF
EQAPSFSEAQGLANIALINNNLTLIQRSNHTQAANGTPLYCLFLCSPT"
16..261
/gene="HLA-D"
/number=2

exon
1..279
/gene="HLA-DPA1"
/number=2

ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTTGGTCTATGCGTCTG 21
|||||
DB 68 CCTGTTGGTCTATGCGTCTG 48

RESULT 14
LOCUS AF015295/c
DEFINITION Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*02013
allele), exon 2 and partial cds.
ACCESSION AF015295

```

```

/product="MHC class II antigen"
/protein_id="AAD42928.1"
/db_xref="GI:5381295"
/translation="HVSYYAAFVOTHRPTGFMFPEFDEQFYVDLKKETVWHLERF
GRAFSFEAQGLANIALINNNLTLIQRSNHTQAANGTPLYCLFLCSPT"
1..279
/gene="HLA-DPA1"
/number=2

ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTTGGTCTATGCGTCTG 21
|||||
DB 48 CCTGTTGGTCTATGCGTCTG 28

RESULT 13
LOCUS HSU87556/c
DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2,
partial cds.
ACCESSION U87556
VERSION U87556.1 GI:2760313
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 286)
Steiner, L., Begovich, A. and Suraj, V.
Direct Submission
JOURNAL Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
COMMENT On Jan 8, 1998 this sequence version replaced gi:1842112.
FEATURES
source
1..286
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
<1..>286
/gene="HLA-D"
<16..>261
/gene="HLA-D"
/note="MHC class II HLA-DPA1 antigen"
/codon_start=3
/protein_id="AAB97110.1"
/db_xref="GI:2795772"
/translation="DHVSTYAAFVOTHRPTGFMFPEFDEQFYVDLKKETVWHLERF
EQAPSFSEAQGLANIALINNNLTLIQRSNHTQAANGTPLYCLFLCSPT"
16..261
/gene="HLA-D"
/number=2

exon
1..279
/gene="HLA-DPA1"
/number=2

ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTTGGTCTATGCGTCTG 21
|||||
DB 68 CCTGTTGGTCTATGCGTCTG 48

RESULT 14
LOCUS AF015295/c
DEFINITION Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*02013
allele), exon 2 and partial cds.
ACCESSION AF015295

```

```

VERSION      AF015295.1  GI:3660653
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 287)
AUTHORS      Steiner, L., Segovitch, A., and Zimmermann, P.
              Submitted (21-JUL-1997) Human Genetics, Roche Molecular Systems,
              Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA
FEATURES     Location/Qualifiers
             source
               1..287
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="6"
               <1..287
               /gene="HLA-DPA1"
               /allele="HLA-DPA1*02013"
               <1..15
               /gene="HLA-DPA1"
               /number=1
               <16..261
               /gene="HLA-DPA1"
               /product="MHC class II antigen"
               <16..261
               /gene="HLA-DPA1"
               /codon_start=3
               /product="MHC class II antigen"
               /protein_id="AAC61669.1"
               /db_xref="GI:3660654"
               /translation="DHVSTYAAFYQTHRPTGEFMFEDEDEQFYVDLDKKEVWHLSE
               FGSAFSPQAQGLANIALNNLNLTIORSNHTQAA"
               16..261
               /gene="HLA-DPA1"
               /number=2
               /number=2
             intron
             mRNA
             CDS
             exon
             intron
ORIGIN
  Query Match      100.0%; Score 21; DB 9; Length 287;
  Best Local Similarity 100.0%; Pred. No. 0.56;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  CCTGTGGTCTATGCGTCTG 21
        |||||
Db      58  CCTGTGGTCTATGCGTCTG 48

RESULT 15
AX237304      294 bp      DNA      linear      PAT 26-SEP-2001
DEFINITION    Sequence 280 from Patent WO0164886.
ACCESSION    AX237304
VERSION      AX237304.1  GI:15796858
KEYWORDS      Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Gaiger, A., Algate, P.A. and Mannion, J.

```

```

TITLE      Compositions and methods for the detection, diagnosis and therapy
            of hematological malignancies
JOURNAL      Patent: WO 0164886-A 280 07-SEP-2001;
            CORIXA CORPORATION (US)
FEATURES     Location/Qualifiers
             source
               1..294
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
  Query Match      100.0%; Score 21; DB 6; Length 294;
  Best Local Similarity 100.0%; Pred. No. 0.56;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  CCTGTGGTCTATGCGTCTG 21
        |||||
Db      87  CCTGTGGTCTATGCGTCTG 107

RESULT 16
AX237554      294 bp      DNA      linear      PAT 26-SEP-2001
LOCUS
DEFINITION    Sequence 530 from Patent WO0164886.
ACCESSION    AX237554
VERSION      AX237554.1  GI:15797108
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Gaiger, A., Algate, P.A. and Mannion, J.
TITLE        Compositions and methods for the detection, diagnosis and therapy
            of hematological malignancies
JOURNAL      Patent: WO 0164886-A 530 07-SEP-2001;
            CORIXA CORPORATION (US)
FEATURES     Location/Qualifiers
             source
               1..294
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
  Query Match      100.0%; Score 21; DB 6; Length 294;
  Best Local Similarity 100.0%; Pred. No. 0.56;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  CCTGTGGTCTATGCGTCTG 21
        |||||
Db      87  CCTGTGGTCTATGCGTCTG 107

RESULT 17
BD058339/c    466 bp      DNA      linear      PAT 27-AUG-2002
LOCUS
DEFINITION    Secrated expressed sequence tags (sESTs).
ACCESSION    BD058339
VERSION      BD058339.1  GI:22603945
KEYWORDS      Zea mays
SOURCE       Zea mays
ORGANISM     Zea mays
REFERENCE    1 (bases 1 to 466)
AUTHORS      Jacobs, K., Mccoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,
            Treacy, M., Spaulding, V. and Agostino, M.J.
TITLE        Secrated expressed sequence tags (sESTs)
JOURNAL      Patent: JP 2001519666-A 194 23-OCT-2001;
            GENETICS INSTITUTE INC
COMMENT      PN JP 2001519666-A/194
            PD 23-OCT-2001

```

```

PF 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TRACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double:
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..466
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /db_xref="taxon:4577"
ORIGIN
Query Match 100.0%; Score 21; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 172 CCCTGTTGGTCTATGCGTCTG 152
|||||
RESULT 18
AX884252/c 476 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 115 from Patent EP1033401.
ACCESSION AX884252
VERSION AX884252.1 GI:40039227
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 115 06-SEP-2000;
Genset (FR)
FEATURES
    source
        1..476
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        95...>475
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAE98366.1"
        /db_xref="GI:40039228"
        /translation="MRPEDRMFHIRAVILRALSLAFLLSLRGAGAIKADHVSTYAAFY
        OTHERPTGFEFPEDEMFYVDLDKETVWHLERFGQAFSPAQGLIAILLNNLN
        TLIQRSNHTQATNDPPEVTVPKPEP"
    sig_peptide
        95..187
        /note="score 10.1 seq SLAFLLSLRGAGA/IX"
ORIGIN
Query Match 100.0%; Score 21; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 247 CCCTGTTGGTCTATGCGTCTG 227
|||||
RESULT 19
BD023862/c 476 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD023862
VERSION BD023862.1 GI:22565085

```

```

KEYWORDS JP 2001269182-A/108.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 476)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 108 02-OCT-2001;
Genset
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/108
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC score 10.1
CC seq SLAFLLSLRGAGA/IK
FH Key Location/Qualifiers
FT CDS 95..475
FT sig_peptide 95..187.
FEATURES
    source
        1..476
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 21; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 247 CCCTGTTGGTCTATGCGTCTG 227
|||||
RESULT 20
AX884251/c 576 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 114 from Patent EP1033401.
ACCESSION AX884251
VERSION AX884251.1 GI:40039225
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 114 06-SEP-2000;
Genset (FR)
FEATURES
    source
        1..576
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        195..>575
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAE98365.1"
        /db_xref="GI:40039226"
        /translation="MRPEDRMFHIRAVILRALSLAFLLSLRGAGAIKADHVSTYAAFY
        OTHERPTGFEFPEDEMFYVDLDKETVWHLERFGQAFSPAQGLIAILLNNLN
        TLIQRSNHTQATNDPPEVTVPKPEP"
    sig_peptide
        195..287
        /note="score 10.1 seq SLAFLLSLRGAGA/IK"

```



```

ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGCTCG 21
|||||
Db 347 CCCTGTTGGTCTATCGGCTCG 327

RESULT 21
BD023861/c
LOCUS      BD023861       576 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION  BD023861
VERSION    BD023861.1 GI:22565084
KEYWORDS  JP 2001269182-A/107.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 576)
AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE     Sequence tag and encoded human protein
JOURNAL   Patent: JP 2001269182-A 107 02-OCT-2001;
GENSET
COMMENT   OS Homo sapiens (human)
           PN JP 2001269182-A/107
           PD 02-OCT-2001
           PF 24-FEB-2000 JP 2000118773
           PR 26-FEB-1999 US 60/122487
           PI JEAN BATUTIST DUMAS MILNE EDWARDS, BIMERIC DUCLAIR, JEAN YVES
           PI JORDAN
           PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
           C12N5/10,
           PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
           G06F15/40
           CC score 10.1
           CC seq SLAFLSLRGAGA/IK
           FH Key      Location/Qualifiers
           FT CDS      195..575
           FT sig_peptide 195..287.
           FT Location/Qualifiers
           1..576
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGCTCG 21
|||||
Db 347 CCCTGTTGGTCTATCGGCTCG 327

RESULT 22
E00485/c
LOCUS      E00485       661 bp      RNA      linear      PAT 29-SEP-1997
DEFINITION DNA sequence of a fragment of pSbalpha-318.
ACCESSION  E00485
VERSION    E00485.1 GI:2168768
KEYWORDS  JP 1985226888-A/3.
SOURCE    unidentified
ORGANISM  unidentified
           unclassified.
REFERENCE  1 (bases 1 to 661)
AUTHORS   Edowado,R.S.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE     NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL   Patent: JP 1985226888-A 3 12-NOV-1985;

ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGCTCG 21
|||||
Db 231 CCCTGTTGGTCTATCGGCTCG 211

RESULT 23
I03086/c
LOCUS      I03086       661 bp      ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 5 from Patent US 4582789.
ACCESSION  I03086
VERSION    I03086.1 GI:268242
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 661)
AUTHORS   Sheldon,E.L.III, Levenson,C.H., Mullis,K.B. and Rapoport,H.
TITLE     Process for labeling nucleic acids using psoralen derivatives
JOURNAL   Patent: US 4582789-A 5 15-APR-1996;
           Cetus Corporation; Emeryville, CA
FEATURES
           source
           1..661
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGCTCG 21
|||||
Db 231 CCCTGTTGGTCTATCGGCTCG 211

RESULT 24
S40633/c
LOCUS      S40633       690 bp      DNA      linear      PRI 06-MAY-1993

```

```

CETUS CORP
OS Human (Homo sapiens)
PN JP 1985226888-A/3
PD 12-NOV-1985
PF 20-MAR-1985 JP 1985054705
PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI
EDOWAADO RUISU SHIERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
KOOREI HAKAADO REBENSON, HENRII RAPOPORUTO
PC C07D519/00, C07H21/04, C12N15/00, C12Q1/68, G01N33/50, G01N33/532,
PC G01N33/58,
PC (C07D519/00, C07D493:04, C07D495:04), (C07D519/00, C07D493:04, PC
C07D493:10);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=Blood;
CC *source: cell_type=B cell;
CC *source: library=cDNA library;
FH Key      Location/Qualifiers
FH misc_feature 1..661
           /note="a fragment derived from pSbalpha-318
           for insertion".
           Location/Qualifiers
           1..661
           /organism="unidentified"
           /mol_type="genomic RNA"
           /db_xref="taxon:32644"

FEATURES
           source
           1..661
           Location/Qualifiers
           1..661
           /organism="unidentified"
           /mol_type="genomic RNA"
           /db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGCTCG 21
|||||
Db 231 CCCTGTTGGTCTATCGGCTCG 211

```

DEFINITION HLA class II: DPAl (DPAl\*0101) [human, Genomic, 690 nt].  
 ACCESSION S40633  
 VERSION S40633.1 GI:1679890  
 KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 690)

AUTHORS Marsh, S.G. and Bodmer, J.G.

TITLE HLA class II nucleotide sequences, 1991

JOURNAL Immunogenetics 33 (5-6), 321-334 (1991)

MEDLINE 91267561

PUBMED 1904836

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gisbseq 40633] from the original journal article.

COMMENT This sequence comes from Figure 15.

On Nov 21, 1996 this sequence version replaced gi:1619630.

Region: HLA class II.

FEATURES

source

1..690 Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

1..690

/gene="DPAl"

/allele="DPAl\*0101"

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGCTCG 21

|||||

Db 60 CCTGTGGTCTATCGCTCG 40

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

/product="MHC class II DP3-alpha"  
 /protein\_id="AA063220.1"

/db\_xref="GI:703089"

/db\_xref="GDB:G00-120-634"

/translation="MRPEDRMFIRKAVILKALSLAFLLSLRGAGAKADHVSYYAFAV

QTRHPTEGMFDEDMFYDLDDKXETVWHLEFGQAFSGAAGGLANLAINNNLN

TLIQSNHQAQNDNDPEVTVEKPEVGLQPTNLICHIDKFFPVLNVTWLCNGELVT

EGVAESLFLPRTDYSFKHSHLYTVSASDFYDCRVEHWGLDQPLKHWAEQFIQMP

ETTEVLCALGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV

36..128

/gene="HLA-DPA1"

/note="G00-120-634"

129..815

/gene="HLA-DPA1"

/product="MHC class II DP3-alpha"

/note="G00-120-634"

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 818;

Best Local Similarity 100.0%; Pred. No. 0.63;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGCTCG 21

|||||

Db 188 CCTGTGGTCTATCGCTCG 168

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

sig\_peptide

mat\_peptide

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 818;

Best Local Similarity 100.0%; Pred. No. 0.63;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGCTCG 21

|||||

Db 188 CCTGTGGTCTATCGCTCG 168

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

sig\_peptide

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature 587...655  
/note="(TM) transmembrane region"  
misc\_feature 656...700  
/note="(CY) cytoplasmatic region"  
polyA\_site 1048  
/note="polyadenylation site"

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 0.65;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 73 CCTGTGGTCTATCGGCTG 53

## RESULT 27

I03088/c I03088 1140 bp ss-DNA linear PAT 21-MAY-1993  
LOCUS  
DEFINITION Sequence 7 from Patent US 4582789.  
ACCESSION I03088  
VERSION I03088.1 GI:268244  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

## REFERENCE

1 (bases 1 to 1140)  
Sheldon,E.L. III, Levenson,C.H., Mullis,K.B. and Rapoport,H.  
TITLE Process for labeling nucleic acids using psoralen derivatives  
JOURNAL Patent: US 4582789-A 7 15-APR-1986;

## FEATURES

source  
1..1140  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1140;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 232 CCTGTGGTCTATCGGCTG 212

## RESULT 28

E00484/c E00484 1201 bp RNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION DNA sequence of pSBalpha-318.  
ACCESSION E00484  
VERSION E00484.1 GI:2168767  
KEYWORDS JP 1985226888-A/2.

## SOURCE

## ORGANISM

## unclassified.

## REFERENCE

1 (bases 1 to 1201)

Edowaado,R.S.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.

NUCLEIC ACID LABELLING SUBSTANCE AND USE

PATENT: JP 1985226888-A 2 12-NOV-1985;

## COMMENT

OS Human {Homo sapiens}  
FN JP 1985226888-A/2  
PD 12-NOV-1985  
PE 20-MAR-1985 JP 1985054705  
PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI  
EOWARDO RUISU SHERIDON ZA SAADO, KERII BANKUSU MIYURISU, PI  
KOOREI HAWAADO REHENSON, HENRII RAPOPORUTO  
PC C07D519/00,C07H21/04,C12N15/00,C-2Q1/68,G01N33/50,G01N33/532,  
PC G01N33/58,  
PC (C07D519/00,C07D493/04,C07D495/04),(C07D519/00,C07D493/04, PC  
C07D493/10);

CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: tissue\_type=Blood;  
CC \*source: cell\_type=B cell;  
CC \*source: library=cDNA library;  
CC \*source: clone=psBalpha-318;  
FH Key Location/Qualifiers  
FH CDS 1..2201  
FT /gene="psBalpha-318".  
FT Location/Qualifiers  
1..1201  
/organism="unidentified"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:32644"

## FEATURES

## source

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 232 CCTGTGGTCTATCGGCTG 212

## RESULT 29

I03006/c I03006 1201 bp ss-DNA linear PAT 21-MAY-1993  
LOCUS  
DEFINITION Sequence 6 from Patent US 4617261.  
ACCESSION I03006  
VERSION I03006.1 GI:268462

## KEYWORDS

## SOURCE

## ORGANISM

## Unknown.

## unclassified.

## REFERENCE

1 (bases 1 to 1201)

Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and

Watson,R.M.

TITLE Process for labeling nucleic acids and hybridization probes

JOURNAL Patent: US 4617261-A 6 14-OCT-1986;

Cetus Corporation; Emeryville, CA

## FEATURES

## source

1..1201  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATCGGCTG 21  
|||||  
Db 232 CCTGTGGTCTATCGGCTG 212

## RESULT 30

I03423/c I03423 1201 bp ss-DNA linear PAT 21-MAY-1993  
LOCUS  
DEFINITION Sequence 6 from Patent US 4822731.  
ACCESSION I03423  
VERSION I03423.1 GI:270023

## KEYWORDS

## SOURCE

## Unknown.

## unclassified.

## REFERENCE

1 (bases 1 to 1201)

Watson,R.M., Sheldon,E.L. III and Sneed,R.M.

TITLE Process for labeling single-stranded nucleic acids and

hybridization probes

JOURNAL Patent: US 4822731-A 6 18-APR-1989;  
 Cetus Corporation; Emeryville, CA  
 FEATURES Location/Qualifiers  
 source 1..1201  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 0.66;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTTGGTCTATCGCTCG 21  
 |||||

Db 232 CCTCTTGGTCTATCGCTCG 212

## RESULT 31

AX552229/c 1259 bp DNA linear PAT 27-NOV-2002  
 LOCUS Sequence 138 from Patent WO0162927.  
 DEFINITION AX552229  
 ACCESSION AX552229.1 GI:25896467  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

## REFERENCE

1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Banville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K., Amshay, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F., Rosen, B.H., Russo, P.D., Spiro, P.A., Bradley, D.L., Chen, A., Cohen, H.J., Daffo, A., Daniels, S.E., Dufour, G.E., Flores, V., Fong, W.I., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S., Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.E., Yu, J.Y., Bratcher, S.R., Chalup, M.S., Danl, C.R. and Hillman, J.L.

## TITLE

Polypeptides and corresponding polynucleotides for diagnostics and therapeutics

JOURNAL Patent: WO 0162927-A 138 30-AUG-2001;

## FEATURES

source Incyte Genomics, Inc. (US)  
 1..1259  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: LI:1169865.1:2000MAY01"

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1259;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTTGGTCTATCGCTCG 21  
 |||||

Db 367 CCTCTTGGTCTATCGCTCG 347

## RESULT 32

HUMHDC02/c 2986 bp DNA linear PRI 07-JAN-1995  
 LOCUS Human MHC class II lymphocyte antigen (DPw4-alpha-1) gene, exons 2-4.  
 DEFINITION

ACCESSION M23904 J02738 M15446

VERSION M23904.1 GI:188385

KEYWORDS cell surface glycoprotein; class II gene; integral membrane protein; major histocompatibility complex.

## SEGMENT

## SOURCE

2 of 3  
 Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 2986)  
 Gustafsson, K., Widmark, E., Jonsson, A.K., Servenius, B., Sachs, D.H.,

Larhammar, D., Rask, L. and Peterson, P.A.

Class II genes of the human major histocompatibility complex. Evolution of the DP region as deduced from nucleotide sequences of the four genes

J. Biol. Chem. 262 (18), 8778-8786 (1987)

87250502

3036829

COMMENT Original source text: Human T-cell DNA, clone p412-1 and clone p2703-1.

FEATURES Location/Qualifiers  
 source 1..2986  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

gene join(M23903.1:563..1095,1..2986)

intron order(M23903.1:663..1095,1..769)

exon 770..1015

intron /gene="DPw4-alpha-1"

exon /number=2

intron 1016..1355

exon /gene="DPw4-alpha-1"

intron /number=2

exon 1356..1637

intron /gene="DPw4-alpha-1"

exon /number=3

intron 1638..1851

exon /gene="DPw4-alpha-1"

intron /number=3

exon 1852..2018

intron /gene="DPw4-alpha-1"

exon /number=4

intron 2019..2986

exon /gene="DPw4-alpha-1"

intron /notes="does not fit consensus"

exon /number=4

ORIGIN About 2.4 kb after segment 1; chromosome 6p21.3.

Query Match 100.0%; Score 21; DB 9; Length 2986;

Best Local Similarity 100.0%; Pred. No. 0.74;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTTGGTCTATCGCTCG 21  
 |||||

Db 822 CCTCTTGGTCTATCGCTCG 802

## RESULT 33

HSHLASBA/c

LOCUS HSHLASBA 14546 bp

DEFINITION Human HLA-SB (DP) alpha gene.

ACCESSION X03100

VERSION X03100.1 GI:32243

KEYWORDS antigen; cell surface glycoprotein; class II antigen; glycoprotein; inverted repeat; Kpn repetitive sequence; major histocompatibility complex; repetitive sequence.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 14546)

AUTHORS Lawrence, S.K., Das, H.K., Pan, J. and Weissman, S.M.

TITLE The genomic organization and nucleotide sequence of the HLA-SB (DP) alpha gene

JOURNAL Nucleic Acids Res. 13 (20), 7515-7528 (1985)

MEDLINE 86041930

PUBMED 2997750

COMMENT Data kindly reviewed (05-MAY-1987) by S.K. Lawrence.

FEATURES Location/Qualifiers

source 1..14546

/organism="Homo sapiens"

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
complement(1..247)
/gene="HLA-SB beta"
prim_transcript
complement(<1..247)
/gene="HLA-SB beta"
complement(1..78)
/gene="HLA-SB beta"
/number=1
complement(79..247)
/gene="HLA-SB beta"
/number=1
complement(<79..178)
/gene="HLA-SB beta"
/codon_start=1
/protein_id="CAA26886.1"
/db_xref="GI:32244"
/db_xref="SPTREMBL:O19686"
/translation="MMVLQVSAAPRTVLTALMLVLTSTVVGGRATP"
complement(323..374)
/number="beta consensus sequence, put. regulatory region"
514..526
/number="imp. inverted repeat a"
2189..2200
/number="imp. inverted repeat a'"
2461..2512
/number="alpha consensus sequence, put. regulatory region"
2571..11200
/gene="HLA-SB"
join(2571..2749,6334..6579,6920..7201,7416..7582,
10872..11200)
/gene="HLA-SB alpha"
join(2571..2749,6334..6579,6920..7201,7416..7582,
10872..11200)
/gene="HLA-SB alpha"
prim_transcript
2571..11200
/gene="HLA-SB"
2571..2749
/gene="HLA-SB"
/number=1
join(2650..2749,6334..6579,6920..7201,7416..7570)
/gene="HLA-SB alpha"
/codon_start=1
/product="class II antigen"
/protein_id="CAA26887.1"
/db_xref="GI:673417"
/db_xref="GOA:P20036"
/db_xref="SWISS-PROT:P20036"
/translation="WRPDRMFHRAVILRALSLAFLILRGAGAIKADHYSTYAFV
QTHRPTGFHMFDEDEMFYDLDRKKTWHLEEFQAFSFEAQGLANIALNNLN
TLIQSNHTQATNDPEVTFVFKPELVQPNLTICHIDKFPFVNLVTLVCLGELVT
EGVAISLFLPRDYDFKHFIYLTVPSEDFYDCRVEHMGDQPLKHEWAQEPQMP
ETETVLCALGIUIGVIIVGTVLIIKLSRGHDPRAQGL"
2650..2742
/gene="HLA-SB"
2650..2749
/gene="HLA-SB"
/number="precursor fragment"
2724..2725
/gene="HLA-SB"
/number="pot. alternate signal sequence splice site"
2750..6333
/gene="HLA-SB"
/number=1
4964..4990
/gene="HLA-SB"
/number="inverted repeat b"
5063..5088
/gene="HLA-SB"
/number="inverted repeat b'"
6334..6579
/gene="HLA-SB"
/number=2

```

```

intron
6580..6919
/gene="HLA-SB"
/number=2
exon
6920..7201
/gene="HLA-SB"
/number=3
intron
7202..7415
/gene="HLA-SB"
/number=3
exon
7416..7582
/gene="HLA-SB"
/number=4
intron
7583..10871
/gene="HLA-SB"
/number=4
misc_feature
8601..9100
/gene="HLA-SB"
/number="sequence homologous to IgC epsilon genes"
8991..9029
/gene="HLA-SB"
/number="inverted repeat C"
10516..10554
/gene="HLA-SB"
/number="inverted repeat C'"
10872..11200
/gene="HLA-SB"
/number=5
repeat_region
12301..12800
/number="Kpn repetitive sequence"
ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTGTTGGTCTATGCGTCTG 21
|||||
Db 6386 CCTGTGTTGGTCTATGCGTCTG 6366
RESULT 34
BX120009 Human DNA sequence from clone DASS-227B13 on chromosome 6, complete
LOCUS sequence.
DEFINITION BX120009
ACCESSION BX120009.10 GI:33504469
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 64380)
Direct Submission
Submitted (07-AUG-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 7, 2003 this sequence version replaced gi:31335527.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate

```

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
DASS-227B13 is from a DNA-arts SSTO human bac library VECTOR:  
pBe1cBAC11.

#### FEATURES

source  
1. .64380  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="6"  
/clone="DASS-227B13"  
/clone\_lib="DNA-arts-BAC.1-SSTO.1"

#### ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 64380;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCGTG 21  
|||||  
Db 2085 CCTGTGTCATGCGTCGTG 2105

#### RESULT 35

AL805913  
LOCUS  
DEFINITION  
Human DNA sequence from clone Xbba-22D21 on chromosome 6, complete sequence.  
AL805913  
VERSION  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Tracey, A.

#### REFERENCE

1 Direct Submission  
Submitted (23-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk  
On Oct 25, 2002 this sequence version replaced gi:22204654.

#### COMMENT

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquerry@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
XXBac-22D21 is from a DNA-arts QSL human bac library VECTOR:  
pBelcBAC11.

#### FEATURES

source  
1. .106728  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="6"  
/clone="XXBac-22D21"  
/clone\_lib="DNA-arts-BAC.1-QBL.1"

#### ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 106728;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCGTG 21  
|||||

Db 95089 CCTGTGTCATGCGTCGTG 95109

#### RESULT 36

AL645931  
LOCUS  
DEFINITION  
Human DNA sequence from clone XXBac-138A21 on chromosome 6, complete sequence.  
AL645931  
VERSION  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Almeida, J.

#### REFERENCE

1 Direct Submission  
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk  
On Mar 21, 2002 this sequence version replaced gi:19031691.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

#### COMMENT

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) XXBac-138A21 is from a CHORI-501 human bac - PGF cell line library VECTOR:  
pTABBAC2.1

This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/MEC.

## FEATURES

Location/Qualifiers  
 1. .124889  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXbac-138A21"  
 /clone\_lib="CHORI-501"

## ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 124899;  
 Best Local Similarity 100.0%; Pred. No 1.2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGGTCTATGCTCTG 21

Db 77770 CCTGTGGTCTATGCTCTG 77790

## RESULT 37

AC011086

## LOCUS

AC011086 181228 bp DNA linear HTG 24-AUG-2002

## DEFINITION

Homo sapiens chromosome 6 clone RP11-93F3 map 6, WORKING DRAFT

SEQUENCE, 12 unordered pieces.

## ACCESSION

AC011086

## VERSION

AC011086.5 GI:10047675

## KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 181228)

## AUTHORS

Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,

Baldwin,J., Barra,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,

Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,

Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,

McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

## Direct Submission

## TITLE

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 181228)

## AUTHORS

Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Barra,N., Bastien,V., Beda,F., Boguslavsky,L.,

Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,

Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,

DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,

FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,

Graham,L., Grand-Pierre,N., Jones,C., Kann,L., Karatas,A., Horton,L.,

Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Lacroque,K.,

Lamazares,R., Lander,E., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,

Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Mitova,T., Mlenga,V.,

Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,

O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,

Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,

Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,

Sougnuez,C., Spencer,E., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,

Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
 Zimmer,A. and Zody,M.

## TITLE

JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 10, 2000 this sequence version replaced gi:7717099.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1322

Center clone name: 93F\_3

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 166135 bases at least Q40

Consensus quality: 173448 bases at least Q30

Consensus quality: 176936 bases at least Q20

Insert size: 174000; agarose-fp

Insert size: 180128; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases.

NOTE: This is a 'working draft' sequence. It currently

consists of 12 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 27639: contig of 27639 bp in length

27640 27739: gap of 100 bp

27740 28970: contig of 1231 bp in length

28971 92294: gap of 100 bp

92294 92294: contig of 53224 bp in length

92295 82394: gap of 100 bp

82394 82394: contig of 4820 bp in length

82395 87215: gap of 100 bp

87215 93029: contig of 5715 bp in length

93029 93129: gap of 100 bp

93129 93130: contig of 8650 bp in length

93130 101780: gap of 100 bp

101780 108114: contig of 6235 bp in length

108114 108115: gap of 100 bp

108115 108215: contig of 7929 bp in length

108215 116143: gap of 100 bp

116143 116243: gap of 100 bp

116243 133838: contig of 17595 bp in length

133838 133938: gap of 100 bp

133938 153417: contig of 19479 bp in length

153417 153517: gap of 100 bp

153517 178444: contig of 24927 bp in length

178444 178544: gap of 100 bp

178544 181228: contig of 2684 bp in length.

181228 Location/Qualifiers

1. .181228

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="6"

/map="6"

/clone="RP11-93F3"

/clone\_lib="RPC1-11 Human Male BAC"

1. .27639

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left"

## misc\_feature

misc_feature		27740..28970		/note="assembly_fragment"					
misc_feature		29071..82294		/note="assembly_fragment"					
misc_feature		82395..87214		/note="assembly_fragment"					
misc_feature		87315..93029		/note="assembly_fragment"					
misc_feature		93130..101779		/note="assembly_fragment"					
misc_feature		101880..108114		/note="assembly_fragment"					
misc_feature		108215..116143		/note="assembly_fragment"					
misc_feature		116244..133838		/note="assembly_fragment"					
misc_feature		133939..153417		/note="assembly_fragment"					
misc_feature		153518..178444		/note="assembly_fragment"					
misc_feature		178545..181228		/note="assembly_fragment"					
		clone_end:77							
		vector_side:right"							
ORIGIN									
Query Match		100.0%; Score 21; DB 2; Length 181228;							
Best Local Similarity		100.0%; Pred. No. 1.2;							
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY		1 CCCTGTTGCTCATCGCTG 21							
Db		64008 CCCTGTTGCTCATCGCTG 64028							
RESULT 38									
AL662824									
LOCUS		AL662824 187964 bp DNA linear PRI 24-APR-2002							
DEFINITION		Human DNA sequence from clone XXbac-22116 on chromosome 6, complete sequence.							
ACCESSION		AL662824.9 GI:20068657							
VERSION		AL662824.9							
KEYWORDS		Hig.							
SOURCE		Homo sapiens (human)							
ORGANISM		Homo sapiens							
REFERENCE		1 Almeida, J.							
AUTHORS		Direct Submission							
TITLE		Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. E-mail enquiries:							
JOURNAL		Cambridge, UK. E-mail enquiries:							
COMMENT		humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Apr 7, 2002 this sequence version replaced gi:20067828. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-22116 is from a CHORI-502 human bac - COX cell line library VECTOR.							

misc_feature		27740..28970		/note="assembly_fragment"					
misc_feature		29071..82294		/note="assembly_fragment"					
misc_feature		82395..87214		/note="assembly_fragment"					
misc_feature		87315..93029		/note="assembly_fragment"					
misc_feature		93130..101779		/note="assembly_fragment"					
misc_feature		101880..108114		/note="assembly_fragment"					
misc_feature		108215..116143		/note="assembly_fragment"					
misc_feature		116244..133838		/note="assembly_fragment"					
misc_feature		133939..153417		/note="assembly_fragment"					
misc_feature		153518..178444		/note="assembly_fragment"					
misc_feature		178545..181228		/note="assembly_fragment"					
		clone_end:77							
		vector_side:right"							
ORIGIN									
Query Match		100.0%; Score 21; DB 2; Length 181228;							
Best Local Similarity		100.0%; Pred. No. 1.2;							
Matches		21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY		1 CCCTGTTGCTCATCGCTG 21							
Db		64008 CCCTGTTGCTCATCGCTG 64028							
RESULT 38									
AL662824									
LOCUS		AL662824 187964 bp DNA linear PRI 24-APR-2002							
DEFINITION		Human DNA sequence from clone XXbac-22116 on chromosome 6, complete sequence.							
ACCESSION		AL662824.9 GI:20068657							
VERSION		AL662824.9							
KEYWORDS		Hig.							
SOURCE		Homo sapiens (human)							
ORGANISM		Homo sapiens							
REFERENCE		1 Almeida, J.							
AUTHORS		Direct Submission							
TITLE		Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. E-mail enquiries: humquery@sanger.ac.uk							
JOURNAL		Cambridge, UK. E-mail enquiries: humquery@sanger.ac.uk							
COMMENT		On Apr 7, 2002 this sequence version replaced gi:20067828. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-22116 is from a CHORI-502 human bac - COX cell line library VECTOR.							





AUTHORS de Groot,N.N.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands

FEATURES  
source  
1..189  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/isolate="Beatrix"  
/db\_xref="taxon:9598"  
/sex="female"  
/cell\_type="transformed B-cells"  
/tissue\_type="blood"  
/dev\_stage="adult"  
<1..>189  
/gene="Patr-DPA1"  
/allele="Patr-DPA1\*0301"  
<1..>189  
/gene="Patr-DPA1"  
/product="MHC class II DPAl antigen"  
<1..>189  
/gene="Patr-DPA1"  
/codon\_start=1  
/product="MHC class II DPAl antigen"  
/protein\_id="AAB92507.1"  
/db\_xref="GI:2583181"  
/translation="FVQTHRPTEGFPEDEMFYVLDKKETVWELDEFGAFSE"  
AQQGLANIALNNMLLI"

ORIGIN  
Query Match 92.4%; Score 19.4; DB 9; Length 189;  
Best Local Similarity 95.3%; Pred. No. 4.7;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
|||||  
Db 27 CCCTGTTGGTCTATGCGTCTG 7

RESULT 45  
AF026696/c

LOCUS AF026696 189 bp DNA linear PRI 24-DEC-1997  
DEFINITION Pongo pygmaeus MHC class II DPAl antigen Popy-DPA1 gene (Popy-DPA1\*0202 allele), partial cds.  
ACCESSION AF026696  
VERSION AF026696.1 GI:2583184  
KEYWORDS  
SOURCE Pongo pygmaeus (orangutan)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.  
REFERENCE 1 (bases 1 to 74)  
AUTHORS Oetting,N. and Bontrouf,R.E.  
TITLE Evolution of the major histocompatibility complex DPAl locus in primates  
JOURNAL Hum. Immunol. 42 (2), 184-187 (1995)  
MEDLINE 95263288  
PubMed 7744622

REFERENCE 2 (bases 1 to 189)  
AUTHORS de Groot,N.N.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands

FEATURES  
source  
1..189  
/organism="Pongo pygmaeus"  
/mol\_type="genomic DNA"  
/isolate="Guchi"  
/db\_xref="taxon:9600"  
/cell\_type="transformed B-cells"  
/tissue\_type="blood"  
/dev\_stage="adult"  
<1..>189

gene

```

/gene="Popy-DPA1"
/allele="Popy-DPA1*0202"
<1..>189
/gene="Popy-DPA1"
/product="MHC class II DPAl antigen"
<1..>189
CDS
/gene="Popy-DPA1"
/codon_start=1
/product="MHC class II DPAl antigen"
/protein_id="AAB92509.1"
/db_xref="GI:2583185"
/translation="FVQTHRPTGEYMFEDFEQFYVDLKKETVWHLLEFGRAFSFE
AQQALADIAILNNLNIMI"
ORIGIN
Query Match          92.4%; Score 19.4; DB 9; Length 189;
Best Local Similarity 95.2%; Pred. No. 4.7;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCTGTTGCTATGCGTCTG 21
    |||||
Db 27 CCCTGTTGCTATGCTGTCG 7
    |||||

RESULT 46
AF026700/c
LOCUS
DEFINITION
Macaca fascicularis clone Maaf*0201 MHC class II DPAl antigen gene,
partial cds.
ACCESSION
AF026700
VERSION
AF026700.1 GI:2583198
KEYWORDS
SOURCE
Macaca arctoides (stump-tailed macaque)
ORGANISM
Macaca arctoides
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 12 to 74)
Ottung,N. and Bontrop,R.E.
Evolution of the major histocompatibility complex DPAl locus in
primates
JOURNAL
Hum. Immunol. 42 (2), 184-187 (1995)
MEDLINE
95263288
PUBMED
7744622
REFERENCE
2 (bases 1 to 189)
de Groot,N.N.
Direct Submission
AUTHORS
Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research
Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands
JOURNAL
Location/Qualifiers
FEATURES
1..189
source
/organism="Macaca arctoides"
/mol_type="genomic DNA"
/isolate="288"
/db_xref="taxon:9540"
/clone="Maaf*0201"
/sex="female"
/cell_type="transformed B-cells"
/tissue_type="blood"
/dev_stage="adult"
<1..>189
/codon_start=1
/product="MHC class II DPAl antigen"
/protein_id="AAB92719.1"
/db_xref="GI:2583199"
/translation="FVQTHRPTGEYMFEDFEQFYVDLKKETVWHLLEFGRAFSFE
AQQALADIAILNNLNIMI"
ORIGIN
Query Match          92.4%; Score 19.4; DB 9; Length 189;
Best Local Similarity 95.2%; Pred. No. 4.7;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCTGTTGCTATGCGTCTG 21
    |||||
Db 27 CCCTGTTGCTATGCTGTCG 7
    |||||

RESULT 48
AF026704/c
LOCUS
DEFINITION
Macaca fascicularis clone Maaf*0201 MHC class II DPAl antigen gene,
partial cds.
ACCESSION
AF026704
VERSION
AF026704.1 GI:2583200
KEYWORDS

```

SOURCE  
 ORGANISM  
 Macaca fascicularis (crab-eating macaque)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 1 (bases 12 to 74)  
 Otting,N. and Bontrop,R.E.  
 Evolution of the major histocompatibility complex DPAL locus in  
 primates  
 Hum. Immunol. 42 (2), 184-187 (1995)  
 95263288  
 774622  
 2 (bases 1 to 189)  
 de Groot,N.N.  
 Direct Submission  
 Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research  
 Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands  
 Location/Qualifiers  
 1..189  
 /organism="Macaca fascicularis"  
 /mol\_type="genomic DNA"  
 /isolate="77"  
 /db\_xref="taxon:9541"  
 /clone="MaFa\*0201"  
 /cell\_type="transformed B-cells"  
 /tissue\_type="blood"  
 /dev\_stage="adult"  
 <1..\_>189  
 /codon\_start=1  
 /product="MHC class II DPAL antigen"  
 /protein\_id="AAB82720.1"  
 /db\_xref="GI:2583201"  
 /translation="FVQTHRTPTCEYMFPEDEGQFYVDLKKETVWLEFFGRAPSE"  
 AAGGLANAIILNNNLNTII"

Query Match 92.4%; Score 19.4; DB 9; Length 189;  
 Best Local Similarity 95.2%; Pred.No.4.7;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCTGTGGTCTATGCGTCG 21  
 Db 27 CCCTGTGGTCTATGTCGTG 7

RESULT 49  
 AF026705/c  
 LOCUS  
 DEFINITION  
 Macaca mulatta clone Mamu\*0201 MHC class II DPAL antigen gene,  
 partial cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Macaca mulatta (rhesus monkey)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 1 (bases 12 to 74)  
 Otting,N. and Bontrop,R.E.  
 Evolution of the major histocompatibility complex DPAL locus in  
 primates  
 Hum. Immunol. 42 (2), 184-187 (1995)  
 95263288  
 774622  
 2 (bases 1 to 189)  
 de Groot,N.N.  
 Direct Submission  
 Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research  
 Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands  
 Location/Qualifiers  
 1..189  
 /organism="Macaca mulatta"

QY 1 CCTGTGTCATGCGTCTG 21  
 Db 27 CCTGTGTCATGCGTCTG 7

RESULT 51  
 AF026707/c  
 LOCUS 189 bp DNA linear PRI 03-NOV-1997  
 DEFINITION Pan troglodytes clone DPAL\*0201 MHC class II DPAL antigen gene, partial cds.  
 ACCESSION AF026707 GI:2583206  
 KEYWORDS Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 REFERENCE 1 (bases 12 to 74)  
 AUTHORS Otting, N. and Bontrop, R.E.  
 TITLE Evolution of the major histocompatibility complex DPAL locus in primates  
 JOURNAL Hum. Immunol. 42 (2), 184-187 (1995)  
 MEDLINE 95263288  
 PUBMED 7744622  
 REFERENCE 2 (bases 1 to 189)  
 AUTHORS de Groot, N.N.  
 JOURNAL Direct Submission  
 TITLE Submitted (25-SEP-1997) Immunobiology, Biomedical Primate Research Centre, Lange Kleiweg 151, Rijswijk, GH 2280, The Netherlands

FEATURES  
 Location/Qualifiers  
 1..189  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /isolate="de Laatste"  
 /db\_xref="taxon:9598"  
 /clone="DPAL\*0201"  
 /sex="male"  
 /cell\_type="transformed B-cells"  
 /tissue\_type="blood"  
 /dev\_stage="adult"  
 <1..>189  
 /codon\_start=1  
 /product="MHC class II DPAL antigen"  
 /protein\_id="AA382723.1"  
 /db\_xref="GI:2583207"  
 /translation="FVQTHRTGTFMFDEDEQFYVLDLKKETVHLEFGRAFSFE AQQGLANIALNNLNLI"

CDS  
 1 CCTGTGTCATGCGTCTG 21  
 27 CCTGTGTCATGCGTCTG 7

ORIGIN  
 Query Match 92.4%; Score 19.4; DB 9; Length 189;  
 Best Local Similarity 95.2%; Pred. No. 4.7;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
 Db 27 CCTGTGTCATGCGTCTG 7

RESULT 52  
 AF529200/c  
 LOCUS 189 bp mRNA linear PRI 01-AUG-2003  
 DEFINITION Aotus nancymae MHC class II antigen (Aona-DPAL) mRNA, Aona-DPAL\*01 allele, partial cds.  
 ACCESSION AF529200  
 VERSION AF529200.1 GI:33327721  
 KEYWORDS Aotus nancymae (Ma's night monkey)  
 SOURCE Aotus nancymae  
 ORGANISM Aotus nancymae  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.  
 REFERENCE 1 (bases 1 to 189)  
 AUTHORS Estupinan, M., Moya, R., Nino, J., Martinez, P., Suarez, C.F. and

Patarroyo, M.E.  
 Preliminary characterization of DPAL locus in the Owl monkey (Aotus nancymae)  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 189)  
 AUTHORS Estupinan, M., Moya, R., Nino, J., Martinez, P., Suarez, C.F. and Patarroyo, M.E.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JUL-2002) Immunologia, Fundacion Instituto de Immunologia de Colombia, FIDIC, Crr 50 # 26-00, Bogota, DC, Colombia

FEATURES  
 Location/Qualifiers  
 1..189  
 /organism="Aotus nancymae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:37293"  
 <1..>189  
 /gene="Aona-DPAL"  
 /allele="Aona-DPAL\*01"  
 <1..>189  
 /gene="Aona-DPAL"  
 /note="putative alpha 1 domain"  
 /codon\_start=1  
 /product="MHC class II antigen"  
 /protein\_id="AAQ09325.1"  
 /db\_xref="GI:33327722"  
 /translation="FVQTHRTGTFMFDEDEQFYVLDLKKETVHLEFGRAFSFE AQQGLANIALNNLNLI"

exon  
 <1..>189  
 /gene="Aona-DPAL"  
 /number=2

ORIGIN  
 Query Match 92.4%; Score 19.4; DB 9; Length 189;  
 Best Local Similarity 95.2%; Pred. No. 4.7;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTGTGTCATGCGTCTG 21  
 Db 27 CCTGTGTCATGCGTCTG 7

RESULT 53  
 HUMHMA1X/c  
 LOCUS 189 bp DNA linear PRI 07-JAN-1995  
 DEFINITION Human MHC class II HLA-DP-alpha-1 gene, exon 2 from cell line AMA1.  
 ACCESSION M83908  
 VERSION M83908.1 GI:187641  
 KEYWORDS antigen presentation; cell surface glycoprotein; class II gene; integral membrane protein; major histocompatibility complex.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 198)  
 AUTHORS Harada, H., Kimura, A., Dong, R.P., Xu, X.P., Bhatia, K. and Sasazuki, T.  
 TITLE Sequencing and population analysis of four novel HLA-DPAL alleles  
 JOURNAL Hum. Immunol. 35 (3), 173-178 (1992)  
 MEDLINE 93186443  
 PUBMED 1363422  
 COMMENT Original source text: Homo sapiens DNA.

FEATURES  
 Location/Qualifiers  
 1..198  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="6p21.3"  
 /cell\_line="EBV transformed cell line AMA1"  
 /gene="HLA-DPAL"

gene  
 1..198  
 /gene="HLA-DPAL"

exon  
 1..198  
 /gene="HLA-DPAL"

```

/Note="G00-120-634"
/number=2

ORIGIN
Query Match          92.4%; Score 19.4; DB 9; Length 198;
Best Local Similarity 95.2%; Pred. No. 4.7;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGGCTCTG 21
    |||
Db 33 CCTGTGGTCTATGGTCTG 13
    |||

RESULT 54
HSDPA1AI/c
LOCUS
DEFINITION H.sapiens MHC class II HLA-DPA1 gene exon 2, cell line AXAL.
ACCESSION X79477
VERSION X79477.1 GI:496618
KEYWORDS MHC class II DPA1 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Rozenmuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,
Marsh,S.G., Bodmer,J.G. and Tilanus,M.G.
Sequencing-based typing reveals new insight in HLA-DPA1
Polymorphism
Tissue Antigens 45 (1), 57-62 (1995)
95242313
MEDLINE
PUBMED 7725312
REFERENCE 2 (bases 1 to 246)
AUTHORS Rozenmuller,E.H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) E.H. Rozenmuller, Diagnostic DNA Laboratory,
University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
Utrecht, NETHERLANDS
Related sequences: S52453 and D14344.
Related sequences: S52453, D14344 and M83906.
FEATURES
source
1..246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/haplotype="DPA1*0201"
/cell_line="DAUDI and AKIBA"
1..226
/gene="MHC class II HLA-DPA1 gene"
1..226
/gene="MHC class II HLA-DPA1 gene"
/number=2

ORIGIN
Query Match          92.4%; Score 19.4; DB 9; Length 246;
Best Local Similarity 95.2%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGGCTCTG 21
    |||
Db 53 CCTGTGGTCTATGGCTCTG 33
    |||

RESULT 56
HSZ48473/c
LOCUS
DEFINITION H.sapiens HLA DPA1 gene for first domain of MHC class 2 molecule,
alpha-chain (allele DPA1*0203).
ACCESSION Z48473
VERSION Z48473.1 GI:1770743
KEYWORDS alpha-chain; MHC class 2 molecule.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muntau,B., Thyse,T., Pirmez,C. and Horstmann,R.D.
A novel DPA1 allele (DPA1*0203) composed of known epitopes
Tissue Antigens 49 (6), 668-669 (1997)
9737898
MEDLINE
PUBMED 9234495
REFERENCE 2 (bases 1 to 252)
AUTHORS Muntau,B.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1995) Birgit Muntau, Molecular Genetics, Bernhard
Nocht Institute for Tropical Medicine, Hamburg,
Bernhard-Nocht-Str.74, 20359 Hamburg, Germany
On Jan 9, 1997 this sequence version replaced gi:683569.
FEATURES
source
1..252
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="Brazilian White"
/db_xref="taxon:9606"
/chromosome="6"

ORIGIN
Query Match          92.4%; Score 19.4; DB 9; Length 226;
Best Local Similarity 95.2%; Pred. No. 4.7;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGGCTCTG 21
    |||
Db 33 CCTGTGGTCTATGGCTCTG 13
    |||

RESULT 55
HS0201X2/c
LOCUS
DEFINITION H.sapiens MHC class II HLA-DPA1*0201 gene (exon 2).
ACCESSION X82394
VERSION X82394.1 GI:572739
KEYWORDS MHC class II HLA DPA1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

/clone="TC 48-3"
/sex="female"
/tissue_type="blood"
/dev_stage="adult"
31..252
/gene="HLA-DPA1"
31..252
/gene="HLA-DPA1"
/product="first domain of MHC class 2 molecule,
alpha-chain"
/note="allele DPA1*0203"
/number=2

Query Match      92.4%; Score 19.4; DB 9; Length 252;
Best Local Similarity 95.2%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

gene
exon

ORIGIN
Query Match      92.4%; Score 19.4; DB 9; Length 252;
Best Local Similarity 95.2%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATGCGTCTG 21
        |||||
Db      60 CCTGTTGGTCTATGGGTCG 40

RESULT 57
AF165160/c
LOCUS      257 bp DNA linear PRI 01-SEP-2000
DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-DPA1) gene,
             HLA-DPA1*0201 variant allele, exon 2 and partial cds.
ACCESSION AF165160
VERSION    1
KEYWORDS   A new HLA-DPA1 allele, DPA1*02016, identified in African-American
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 257)
AUTHORS   McDaniel,D.O., Nguyen,C. and McDaniel,L.S.
TITLE     A new HLA-DPA1 allele, DPA1*02016, identified in African-American
JOURNAL   Tissue Antigens 56 (2), 197-198 (2000)
MEDLINE   20470607
PUBMED    11019928
REFERENCE  2 (bases 1 to 257)
AUTHORS   McDaniel,D.O., Nguyen,C. and McDaniel,L.S.
TITLE     Direct Submission
JOURNAL   Submitted (03-JUL-1999) Surgery, University of Mississippi Medical
Center, 2500 North State Street, Jackson, MS 39216, USA

FEATURES             source
1..257
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_type="peripheral blood cell"
<1..>257
/allele="HLA-DPA1"
/allele="HLA-DPA1*0201 variant"
<1..>257
/gene="HLA-DPA1"
/gene="HLA-DPA1"
<1..>257
/gene="HLA-DPA1"
/gene="HLA-DPA1"
/codon_start=1
/product="MHC class II HLA-DPA1 antigen"
/protein_id="AAD47826.1"
/db_xref="GI:5713148"
/translation="IKADHVSTYAAVQTHRTGTFMFEDDEQFYVLDKKTWVH
LEEFGRATSFQAQGLANLAINNNTLIQRSHNTQAANDP"
1..257
/gene="HLA-DPA1"
/number=2

exon

ORIGIN
Query Match      92.4%; Score 19.4; DB 9; Length 258;
Best Local Similarity 95.2%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATGCGTCTG 21
        |||||
Db      60 CCTGTTGGTCTATGGGTCG 40

RESULT 58
AF118120/c
LOCUS      258 bp DNA linear PRI 24-MAR-1999
DEFINITION Homo sapiens MHC class II antigen DP alpha 1 subunit HLA-DPA1 gene
             (HLA-DPA1*02013 allele), exon 2 and partial sequence.
ACCESSION AF118120
VERSION    1
KEYWORDS   A novel HLA-DPA1 variant DPA1*02013 found in African-American
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 258)
AUTHORS   McDaniel,D.O., Nguyen,C. and McDaniel,L.S.
TITLE     A novel HLA-DPA1 variant DPA1*02013 found in African-American
JOURNAL   Unpublished
MEDLINE   2 (bases 1 to 258)
AUTHORS   McDaniel,D.O., Nguyen,C. and McDaniel,L.S.
TITLE     Direct Submission
JOURNAL   Submitted (15-JUL-1998) Surgery, UMMC, 2500 North State Street,
Clinical Science Bldg., Jackson, MS 39216, USA
COMMENT    NCBI staff are still waiting for submitters to provide appropriate
coding region information.
FEATURES             Location/Qualifiers
1..258
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_type="peripheral blood mononuclear cells"
<1..>258
/allele="HLA-DPA1"
/allele="HLA-DPA1*02013"
<1..>258
/gene="HLA-DPA1"
/number=2

gene
exon

ORIGIN
Query Match      92.4%; Score 19.4; DB 9; Length 258;
Best Local Similarity 95.2%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CCTGTTGGTCTATGCGTCTG 21
        |||||
Db      60 CCTGTTGGTCTATGGGTCG 40

RESULT 59
AF092049/c
LOCUS      265 bp DNA linear PRI 01-OCT-1999
DEFINITION Homo sapiens isolate 904 MHC class II antigen (HLA-DPA1) gene,
             partial cds.
ACCESSION AF092049
VERSION    1
KEYWORDS   AF092049.1 GI:6002596
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 265)
AUTHORS   Varney,M.D., Gavrilidis,A. and Abbott,W.

```

```

TITLE      DPAl Polymorphism in Polynesians
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 265)
AUTHORS    Varney,X.D., Gavrilidis,A. and Abbott,W.
TITLE      Direct Submission
JOURNAL    Submitted (15-SEP-1998) Tissue Typing, Royal Melbourne Hospital,
           Grattan Street, Parkville, VIC 3050, Australia
FEATURES   Location/Qualifiers
            source      1..265
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /isolate="904"
                        /db_xref="taxon:9606"
gene       <1..>265
           /chromosome="6"
mRNA       <1..>265
           /gene="HLA-DPA1"
           <1..>265
           /gene="HLA-DPA1"
           <1..>265
           /product="MHC class II antigen"
           /codon_start=1
           /protein_id="AAC64233.1"
           /db_xref="GI:3757798"
           /translation="TFCRVFLYFLYAADHVSTYAMFVQTHRPTGFMFEFDEDEMPYV
           DLDKRTVWHLLEEFQGFAPFQAQGLANIALINNNLTIQRSNHTQATNGTGYLCLF
           LCSPTG"
ORIGIN
Query Match      92.4%; Score 19.4; DB 9; Length 326;
Best Local Similarity 95.2%; Pred. No. 5;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CCTCTTGGTCTATGCGTCTG 21
        |||||
Db      90  CCTCTTGGTCTATGCGTCTG 70

Search completed: April 20, 2004, 09:33:55
Job time : 639.447 secs

TITLE      DPAl Polymorphism in Polynesians
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 265)
AUTHORS    Varney,X.D., Gavrilidis,A. and Abbott,W.
TITLE      Direct Submission
JOURNAL    Submitted (15-SEP-1998) Tissue Typing, Royal Melbourne Hospital,
           Grattan Street, Parkville, VIC 3050, Australia
FEATURES   Location/Qualifiers
            source      1..265
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /isolate="904"
                        /db_xref="taxon:9606"
gene       <1..>265
           /chromosome="6"
mRNA       <1..>265
           /gene="HLA-DPA1"
           <1..>265
           /gene="HLA-DPA1"
           <1..>265
           /product="MHC class II antigen"
           /codon_start=1
           /protein_id="AAF0051.1"
           /db_xref="GI:6002597"
           /translation="VSTYAMFVQTHRPTGFMFEFDEDEQFVVDLDKRTVWHLLEEFQ
           RAFSFEAQGLANIALINNNLTIQRSNHTQAANGTAYLCLFL"
ORIGIN
Query Match      92.4%; Score 19.4; DB 9; Length 265;
Best Local Similarity 95.2%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CCTCTTGGTCTATGCGTCTG 21
        |||||
Db      47  CCTCTTGGTCTATGCGTCTG 27

RESULT 60
AF013767/c
LOCUS      AF013767      326 bp      DNA      linear      PRI 16-OCT-1998
DEFINITION Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*0302
            allele), partial cds.
ACCESSION  AF013767 U94838
VERSION    AF013767.1 GI:3660651
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 326)
AUTHORS    Steiner,L.L., Cavalli,A., Zimmerman,P.A., Boatín,B.A.,
            Titanji,V.P., Bradley,J.Z., Lucius,R., Nutman,T.B. and
            Begovich,A.B.
           Three new DP alleles identified in sub-Saharan Africa: DPB1*7401,
           DPA1*02013, and DPA1*0302
JOURNAL    Tissue Antigens 51 (6), 653-657 (1998)
MEDLINE    98357732
PUBMED     9694359
REFERENCE   2 (bases 1 to 326)
AUTHORS    Steiner,L., Begovich,A. and Zimmerman,P.
TITLE      Direct Submission
JOURNAL    Submitted (14-JUL-1997) Human Genetics, Roche Molecular Systems,
           1145 Atlantic Ave., Alameda, CA 94501, USA
COMMENT    On Oct 16, 1998 this sequence version replaced gi:2865247.
FEATURES   Location/Qualifiers
            source      1..326
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 134.771 Seconds  
(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819B-40

Perfect score: 21  
Sequence: 1 cccgtgtggtctatgcgtctg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

N: Geneseq\_29Jan04.\*

1: geneseqn1980s.\*

2: geneseqn1990s.\*

3: geneseqn2000s.\*

4: geneseqn2001as.\*

5: geneseqn2001bs.\*

6: geneseqn2002s.\*

7: geneseqn2003as.\*

8: geneseqn2003bs.\*

9: geneseqn2003cs.\*

10: geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	6	AAL48204 Human HLA
2	21	100.0	255	6	AAL48219 Human HLA
3	21	100.0	267	4	AAK54418 Human hae
4	21	100.0	267	4	AAK54603 Human hae
5	21	100.0	272	4	AAK54317 Human hae
6	21	100.0	294	4	AAK54555 Human hae
7	21	100.0	294	4	AAK54805 Human hae
8	21	100.0	362	3	AAK43013 Human hae
9	21	100.0	410	8	ACH49248 Human leu
10	21	100.0	436	3	AAK43818 Human sec
11	21	100.0	456	2	AAV86216 Human sec
12	21	100.0	476	3	AAC00117 Human sec
13	21	100.0	490	8	ACH49328 Human leu
14	21	100.0	576	3	AAC00116 Human sec
15	21	100.0	1202	2	AAC25060 Human sec
16	21	100.0	1259	4	AAK31123 Human dia
17	21	100.0	1348	3	AAT18332 Lung canc
18	21	100.0	14646	6	ABK64796 Human ben
19	20	95.2	20	6	AAL48203 Human HLA
20	19.4	92.4	396	2	AAV86130 EST clone
21	19	90.5	4233	3	AAZ90198 Rat mdrlb
22	19	90.5	4233	4	AAZ27498 Rat mdrlb
23	19	90.5	4254	6	ABK63517 Rat seque

24	19	90.5	4254	7	ABT41782	Abt41782 Toxicity
25	17.8	84.8	110000	4	RAI99682_03	Continuation (4 of
26	17.8	84.8	110000	4	RAI99683_03	Continuation (4 of
27	17.4	82.9	96596	8	ADA02564	Human RAS
28	17.4	82.9	96596	9	ADB72302	Human RAS
29	17.4	82.9	240000	7	ACD13446	Human DNA
30	16.8	80.0	267	2	AAV86047	EST clone
31	16.4	78.1	4189	3	AAZ49334	Murine mu
32	16.4	78.1	4189	3	AAZ49334	Mouse SCR
33	16.4	78.1	4298	9	ADD29583	Mouse tum
34	16.4	78.1	4313	2	AAQ38950	Mouse mul
35	16.2	77.1	396	4	AAI82378	Human pol
36	16.2	77.1	423	5	ABA11047	Human ner
37	16.2	77.1	1847	5	ABA11047	Human ner
38	16.2	77.1	7785	2	AAV81446	Pig p105
39	16.2	77.1	11348	4	AAK80405	Human imm
40	16.2	77.1	11350	4	AAK80407	Human imm
41	16.2	77.1	12642	7	ABT31904	Human bre
42	16.2	77.1	12680	6	ABT07741	Breast ca
43	16.2	77.1	12680	9	ABT07518	Prostate
44	16.2	77.1	12716	6	ABN59934	Novel hum
45	16.2	77.1	17069	4	AAK83892	Human imm
46	16.2	77.1	17069	4	AAK80906	Human imm
47	16.2	77.1	53552	5	AAI31365	Genomic D
48	16	76.2	279	2	AAZ32177	Human coa
49	16	76.2	148567	7	ABS55500	Gene enco
50	16	76.2	148567	8	ACA82841	Human kin
51	15.8	75.2	65	6	ABN52705	Mouse spl
52	15.8	75.2	271	6	ABS51596	Human cdn
53	15.8	75.2	612	6	ABN63360	Human can
54	15.8	75.2	1183	6	ABL55410	Human l1
55	15.8	75.2	26410	4	AAK70623	Human imm
56	15.8	75.2	40116	7	ABZ26080	Mouse DNA
57	15.8	75.2	96599	9	ADA02981	Mouse Map
58	15.8	75.2	96599	9	ADB72719	Mouse Map
59	15.8	75.2	96599	9	ADC85461	Mouse Map
60	15.8	75.2	110000	7	ADA53224_2	Continuation (3 of
61	15.8	75.2	112132	6	ABK90888	Human ATP
62	15.8	75.2	148834	6	ABK83570	Human cdn
63	15.8	75.2	322101	9	ADA58431	Human FAO
64	15.4	73.3	565	5	ABV55407	Human pro
65	15.4	73.3	1068	8	ADA31143	DNA encod
66	15.4	73.3	1353	7	ACA23947	Prokaryot
67	15.4	73.3	27150	4	AAZ02701	Human gly
68	15.4	73.3	54391	7	ACR42745_3	Continuation (4 of
69	15.4	73.3	172637	6	ABN83124	Human vol
70	15.4	73.3	237961	6	ABQ80552	Human Can
71	15.2	72.4	140	4	ABA66598	Human foe
72	15.2	72.4	140	4	AAK40756	Human bon
73	15.2	72.4	140	4	AAK15025	Human bra
74	15.2	72.4	140	4	ABS40330	Human liv
75	15.2	72.4	295	3	AAZ97695	Bridge-1
76	15.2	72.4	395	6	ABL66554	Lung canc
77	15.2	72.4	439	6	ABN95367	Gene #186
78	15.2	72.4	446	6	ABL78941	Human ova
79	15.2	72.4	454	4	AAI08677	Human bre
80	15.2	72.4	477	4	ABA54027	Human foe
81	15.2	72.4	477	4	AAK27744	Human bon
82	15.2	72.4	477	4	AAK02300	Human bra
83	15.2	72.4	477	4	ABS27322	Human liv
84	15.2	72.4	519	7	ABT42436	Toxicity
85	15.2	72.4	519	9	ADB53851	Primary r
86	15.2	72.4	547	4	AAK78196	Human imm
87	15.2	72.4	547	4	AAK78196	Human imm
88	15.2	72.4	547	4	AAK78197	Human imm
89	15.2	72.4	630	6	ABQ54398	Human ova
90	15.2	72.4	666	4	AAZ71573	Corynebac
91	15.2	72.4	672	4	AAZ71574	Corynebac
92	15.2	72.4	912	4	AAH98478	Human EST
93	15.2	72.4	933	3	AAH98506	Human EST
94	15.2	72.4	1050	3	AAH98506	Human ves
95	15.2	72.4	1148	3	AAZ47965	Arabidops
96	15.2	72.4	1178	2	AAV69563	Soil deri

```

97 15.2 72.4 1356 3 AA253678 Aar53678 Neisseria
98 15.2 72.4 1449 5 AAS65954 Aas6594 DNA encod
99 15.2 72.4 1438 5 AD309924 Ade09924 Novel DNA
100 15.2 72.4 1554 5 AAH67693 Aah67693 C glutam
101 15.2 72.4 1691 3 AAF59769 Aac59769 Human sec
102 15.2 72.4 2122 4 AAF26582 Aaf26582 DNA encod
103 15.2 72.4 2122 7 ABZ73292 Abz73292 Secreted
104 15.2 72.4 2122 9 ADC20072 Adc20072 Human sec
105 15.2 72.4 2259 5 AAS85920 Aas85920 DNA encod
106 15.2 72.4 2270 9 ADE54693 Ade54693 Human gen
107 15.2 72.4 2338 6 AAL49314 Aal49314 Rat Glut1
108 15.2 72.4 2338 6 AAL49313 Aal49313 Rat Glut1
109 15.2 72.4 2338 6 AAL49315 Aal49315 Rat Glut1
110 15.2 72.4 2571 9 AD53698 Ad53698 Primary r
111 15.2 72.4 3175 4 AAI99170 Aai99170 Human exc
112 15.2 72.4 3175 4 AAS32858 Aas32858 Human gen
113 15.2 72.4 3175 5 AAI63520 Aai63520 Human kid
114 15.2 72.4 3202 7 ABT18064 Abt18064 Aspergill
115 15.2 72.4 3203 7 ABT19878 Abt19878 Aspergill
116 15.2 72.4 3465 5 ABV25341 Abv25341 Human pro
117 15.2 72.4 6064 4 AAK78298 Aak78298 Human imm
118 15.2 72.4 6064 4 AAK78297 Aak78297 Human imm
119 15.2 72.4 17310 4 AAK83881 Aak83881 Human imm
120 15.2 72.4 17897 9 ADB53995 Adb53995 DRG1 geno
121 15.2 72.4 26379 3 AAZ88922 Aaz88922 Human wol
122 15.2 72.4 29634 4 AAS59539 Aas5939 Propionib
123 15.2 72.4 29634 7 ACF64468 Acf64468 Propionib
124 15.2 72.4 44567 7 AD56117 Ad56117 Human BAT
125 15.2 72.4 44567 8 ADA02479 Adac02479 Human BAT
126 15.2 72.4 44567 9 ADB72218 Adb72218 Human BAT
127 15.2 72.4 67212 3 AAA08954 Aaa08954 WFS1 vari
128 15.2 72.4 96592 8 ADA02678 Ada02678 Human TOP
129 15.2 72.4 96592 9 ADB72416 Adb72416 Human TIO
130 15.2 72.4 123219 4 AAB88703 Aab8703 Human DNA
131 15.2 72.4 123219 4 AAB88703 Aab8703 Human DNA
132 15.2 72.4 34980 5 ABQ75562 Abq75562 Human rel
133 15.2 71.4 474 4 ABL27511 Abl27511 Drosophil
134 15.2 71.4 2590 4 ABL27510 Abl27510 Drosophil
135 15.2 71.4 3188 4 ABL27500 Abl27500 Drosophil
136 15.2 71.4 3448 4 ABL27502 Abl27502 Drosophil
137 15.2 71.4 3600 4 ABL18800 Abl18800 Drosophil
138 15.2 71.4 110000 7 AAD53223.0 Aad53223 Human chr
139 14.8 70.5 157 4 AAK68741 Aak68741 Human imm
140 14.8 70.5 191 7 ACA14652 Aca14652 Prokaryot
141 14.8 70.5 223 7 ACA15030 Aca15030 Prokaryot
142 14.8 70.5 239 7 ACA14760 Aca14760 Prokaryot
143 14.8 70.5 239 7 ACA14846 Aca14846 Prokaryot
144 14.8 70.5 248 7 ACA14871 Aca14871 Prokaryot
145 14.8 70.5 278 7 ACA14875 Aca14875 Prokaryot
146 14.8 70.5 303 7 ACA14728 Aca14728 Prokaryot
147 14.8 70.5 325 7 ACA14488 Aca14488 Prokaryot
148 14.8 70.5 354 7 ACA14724 Aca14724 Prokaryot
149 14.8 70.5 355 7 ACA14872 Aca14872 Prokaryot
150 14.8 70.5 456 7 ACA14500 Aca14500 Prokaryot
151 14.8 70.5 457 7 ACA15187 Aca15187 Prokaryot
152 14.8 70.5 570 7 ACA15253 Aca15253 Prokaryot
153 14.8 70.5 754 5 AAS86696 Aas86696 DNA encod
154 14.8 70.5 832 5 AAS86701 Aas86701 DNA encod
155 14.8 70.5 1062 4 AAH34142 Aah34142 Human col
156 14.8 70.5 1080 7 ACA52621 Aca52621 Prokaryot
157 14.8 70.5 1097 7 ACA52621 Aca52621 Prokaryot
158 14.8 70.5 1274 4 AAF63382 Aaf63382 Arf promo
159 14.8 70.5 1284 7 ACA49506 Aca49506 Prokaryot
160 14.8 70.5 1329 7 ACA32059 Aca32059 Prokaryot
161 14.8 70.5 1332 3 AAF65960 Aaf65960 E. coli p
162 14.8 70.5 1332 4 AAS58363 Aas58363 Salmonell
163 14.8 70.5 1332 4 AAS2605 Aas2605 E. coli D
164 14.8 70.5 1332 4 AAH84515 Aah84515 E. coli g
165 14.8 70.5 1332 5 AAH1296 Aah1296 Escherich
166 14.8 70.5 1332 7 ACA18726 Aca18726 Prokaryot
167 14.8 70.5 1332 7 ACA54227 Aca54227 Prokaryot
168 14.8 70.5 1332 7 ACA51678 Aca51678 Prokaryot
169 14.8 70.5 1369 3 AAC40089 Aac40089 Arabidops

```

## ALIGNMENTS

```

RESULT 1
AAL48204
ID AAL48204 standard; DNA; 21 BP.
AC AAL48204;
XX
XX
DT 01-OCT-2002 (first entry)
XX
XX Human HLA DPAL locus polymorphism address tag sequence #3.
DE
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX flow cytometry; human; DPAL; DPBI; address tag; ss.
KW Homo sapiens.
OS
XX WO200194639-A1.
XX
XX 13-DEC-2001.
XX
XX 07-JUN-2001; 2001WO-US018590.
XX
XX 08-JUN-2000; 2000US-0210759P.
XX (REGC ) UNIV CALIFORNIA.
XX
XX White PS, Torrey DC;
XX
XX WPI; 2002-566450/60.
XX
XX Identifying sequences useful as address/capture tags for flow cytometry
XX based minisequencing, by generating tag sequences and rejecting sequences
XX based on certain parameters e.g. sequences which form stable hairpins.
XX
XX Disclosure; Page 14; 35pp; English.
XX
XX The present invention relates to a method of identifying sequences useful
XX as address/capture tags, involving rejecting sequences having common sub-
XX sequences with a sub-sequence length greater than specified number of
XX bases, and sequences which can form stable hairpins and stable dimers
XX from a sample of oligonucleotides, and selecting those sequences in the
XX sample that would hybridise to their respective complements with a high
XX degree of specificity. The method is useful for identifying a set of
XX sequences useful as address/capture tags which can be used for
XX multiplexed single nucleotide polymorphism (SNP) scoring in a flow
XX cytometry assay. The present sequence is an address tag described in the
XX exemplification of the invention
XX
XX Sequence 21 BP; 1 A; 6 C; 6 G; 8 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 21; DB 6; Length 21;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTGGTCTATGGCTCTG 21

Db 1 CCCTGTGGTCTATGGCTCTG 21

```

RESULT 2
AA48219/c
ID AAL49219 standard; DNA; 255 BP.
XX
XX AAL48219;
AC
DT 01-OCT-2002 (first entry)
DE
DE Human HLA DPAL exon 2 sequence.
XX
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
KW flow cytometry; human; DPAL; DPB1; Gene; ds.
XX
XX Homo sapiens.
XX
XX W0200194639-A1.
PN
XX 13-DEC-2001.
PD
XX
XX 07-JUN-2001; 2001WO-US018590.
PF
XX
XX 08-JUN-2000; 2000US-0210759P.
PR
XX (REGC ) UNIV CALIFORNIA.
PA
XX White PS, Torney DC;
PI
XX WPI; 2002-566450/60.
DR
XX
XX Identifying sequences useful as address/capture tags for flow cytometry
PT based minisequencing, by generating tag sequences and rejecting sequences
PT based on certain parameters e.g. sequences which form stable hairpins.
XX
XX Disclosure; Fig 4; 35pp; English.
XX
XX The present invention relates to a method of identifying sequences useful
CC as address/capture tags, involving rejecting sequences having common sub-
CC sequences with a sub-sequence length greater than specified number of
CC bases, and sequences which can form stable hairpins and stable dimers
CC from a sample of oligonucleotides, and selecting those sequences in the
CC sample that would hybridise to their respective complements with a high
CC degree of specificity. The method is useful for identifying a set of
CC sequences useful as address/capture tags which can be used for
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
CC cytometry assay. The present sequence is a fragment of the human DPAL
CC gene described in the exemplification of the invention
XX
XX Sequence 255 BP; 67 A; 57 C; 66 G; 65 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 60 CCCGTGTTGGTCTATGCGTCTG 40

RESULT 3
AAK54418/c
ID AAK54418 standard; cDNA; 267 BP.
XX
XX AAK54418;
AC
XX
XX 13-NOV-2001 (first entry)
DT
DE Human haematological malignancy-related antigen coding sequence #143.
DE
DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX

```

```

OS Homo sapiens.
XX
XX W0200164886-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 01-MAR-2001; 2001WO-US007272.
PF
XX
XX 01-MAR-2000; 2000US-0186126P.
PR
XX 17-MAR-2000; 2000US-0190479P.
PR
XX 27-APR-2000; 2000US-0200545P.
PR
XX 28-APR-2000; 2000US-0200303P.
PR
XX 28-APR-2000; 2000US-0200779P.
PR
XX 01-MAY-2000; 2000US-0200999P.
PR
XX 04-MAY-2000; 2000US-0202084P.
PR
XX 22-MAY-2000; 2000US-0206201P.
PR
XX 14-JUN-2000; 2000US-0218950P.
PR
XX 03-AUG-2000; 2000US-0222903P.
PR
XX 04-AUG-2000; 2000US-0223416P.
PR
XX 07-AUG-2000; 2000US-0223378P.
XX
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J;
PI
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 371; 1252pp; English.
PS
XX
XX The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 181 CCCGTGTTGGTCTATGCGTCTG 161

RESULT 4
AAK54603/c
ID AAK54603 standard; cDNA; 267 BP.
XX
XX AAK54603;
AC
XX
XX 13-NOV-2001 (first entry)
DT
DE Human haematological malignancy-related antigen coding sequence #328.
DE
DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX W0200164886-A2.
PN

```

```

XX 07-SEP-2001.
PD
XX
XX 01-MAR-2001; 2001WO-US007272.
PF
XX
XX 01-MAR-2000; 2000US-0186126P.
PR
XX
XX 17-MAR-2000; 2000US-0190479P.
PR
XX
XX 27-APR-2000; 2000US-0200545P.
PR
XX
XX 28-APR-2000; 2000US-0200303P.
PR
XX
XX 01-MAY-2000; 2000US-0200999P.
PR
XX
XX 04-MAY-2000; 2000US-020084P.
PR
XX
XX 14-JUL-2000; 2000US-0218950P.
PR
XX
XX 03-AUG-2000; 2000US-0222903P.
PR
XX
XX 07-AUG-2000; 2000US-0223378P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Gaiger A, Algate PA, Mannion J;
PI
XX
XX WPI; 2001-514842/56.
DR
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 418; 1252pp; English.
PS
XX
XX The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of hematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 181 CCCTGTTGGTCTATGCGTCTG 161

RESULT 5
AAK54317/c
ID AAK54317 standard; cDNA; 272 BP.
XX
XX AC AAK54317;
XX
XX 13-NOV-2001 (first entry)
DT
XX
XX Human haematological malignancy-related antigen coding sequence #42.
DE
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200164886-A2.
PN
XX
XX 07-SEP-2001.
XX
XX

```

```

PF 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
PR
XX
XX 17-MAR-2000; 2000US-0190479P.
PR
XX
XX 27-APR-2000; 2000US-0200545P.
PR
XX
XX 28-APR-2000; 2000US-0200303P.
PR
XX
XX 01-MAY-2000; 2000US-0200999P.
PR
XX
XX 04-MAY-2000; 2000US-020084P.
PR
XX
XX 14-JUL-2000; 2000US-0218950P.
PR
XX
XX 03-AUG-2000; 2000US-0222903P.
PR
XX
XX 07-AUG-2000; 2000US-0223378P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Gaiger A, Algate PA, Mannion J;
PI
XX
XX WPI; 2001-514842/56.
DR
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 345; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma
XX
XX Sequence 272 BP; 67 A; 60 C; 75 G; 70 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 186 CCCTGTTGGTCTATGCGTCTG 166

RESULT 6
AAK54555
ID AAK54555 standard; cDNA; 294 BP.
XX
XX AC AAK54555;
XX
XX 13-NOV-2001 (first entry)
DT
XX
XX Human haematological malignancy-related antigen coding sequence #280.
DE
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200164886-A2.
PN
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
PR

```

```

PR 17-MAR-2000; 2000US-0190479P.
PR 27-APR-2000; 2000US-0200545P.
PR 28-APR-2000; 2000US-0200303P.
PR 28-APR-2000; 2000US-0200779P.
PR 01-MAY-2000; 2000US-0200999P.
PR 04-MAY-2000; 2000US-0202084P.
PR 22-MAY-2000; 2000US-0218950P.
PR 14-JUL-2000; 2000US-0206201P.
PR 03-AUG-2000; 2000US-0222903P.
PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
PA
XX
XX
PI Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 46; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of hematological malignancies. The
CC present sequence is the coding sequence of a human hematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of hematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma
XX
XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTTGGTCTATCGGTCGTG 21
DB 87 CCTCTTGGTCTATCGGTCGTG 107
RESULT 7
AAK54805
ID AAK54805 standard; cDNA; 294 BP.
AC AAK54805;
XX
XX 13-NOV-2001 (first entry)
DE
DE Human hematological malignancy-related antigen coding sequence #530.
KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW hematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX Homo sapiens.
OS
XX WO200164886-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 01-MAR-2001; 2001WO-US007272.
PF
XX
XX 01-MAR-2000; 2000US-0186126P.
PR
XX 17-MAR-2000; 2000US-0190479P.
PR
XX 27-APR-2000; 2000US-0200545P.
PR
XX 28-APR-2000; 2000US-0200303P.

```

```

PR 28-APR-2000; 2000US-0200779P.
PR 01-MAY-2000; 2000US-0200999P.
PR 04-MAY-2000; 2000US-0202084P.
PR 22-MAY-2000; 2000US-0206201P.
PR 14-JUL-2000; 2000US-0218950P.
PR 03-AUG-2000; 2000US-0222903P.
PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
PA
XX
XX
PI Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 46; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of hematological malignancies. The
CC present sequence is the coding sequence of a human hematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of hematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma
XX
XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 21; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTTGGTCTATCGGTCGTG 21
DB 87 CCTCTTGGTCTATCGGTCGTG 107
RESULT 8
AAA43013/c
ID AAA43013 standard; cDNA; 362 BP.
XX
XX AAA43013;
AC
XX
XX 21-AUG-2000 (first entry)
DT
XX
XX Human secreted expressed sequence tag SEQ ID NO:1753.
DE
XX
XX Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
XX expressed sequence tag; EST; probe; chemotactic; proliferative;
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antidiastmatic; vulnerary; antiparkinsonian;
XX antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
XX autoimmune disorder; multiple sclerosis; allergic condition;
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
XX central nervous system disorder; Alzheimer's disease; stroke;
XX Parkinson's disease; Huntington's disease; coagulation disorder;
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
XX infection; depression; psoriasis; ss.
XX
XX Homo sapiens.
OS
XX WO200021990-A1.
PN
XX

```

PD 20-APR-2000.  
XX  
XX  
PF 15-OCT-1999; 99WO-US024205.  
XX  
XX  
PR 15-OCT-1998; 98US-0104435P.  
XX  
XX  
PA (GEMY ) GENETICS INST INC.  
XX  
XX  
PI Jacobs K, McCoy JV, Lavallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M;  
XX  
XX  
DR WPI; 2000-317937/27.  
XX  
XX  
PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (sests), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders.  
XX  
XX  
PS Claim 1; Page 526; 618pp; English.  
XX  
XX  
CC AAA41261 to AAA43419 represent specifically claimed secreted expressed  
CC sequence tags (sests), isolated from human, mouse, xenopus and rat tissue  
CC sources. The sestS can have a range of activities depending on the  
CC tissues they were isolated from. The activities include: chemotactic;  
CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antitumor;  
CC osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;  
CC cerebroprotective; anticonvulsant; and antidepressant. The sestS can be  
CC used for gene therapy and in vaccines. The sestS are useful as probes for  
CC the identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the sestS. Proteins encoded by the sestS  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
CC in the exemplification of the present invention  
XX  
XX  
SQ Sequence 362 BP; 86 A; 86 C; 99 G; 91 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 21; DB 3; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CCTGTGCTCTATGCGTCTG 21  
DB 185 CCTGTGCTCTATGCGTCTG 165  
  
RESULT 9  
ACH49248/c  
ID ACH49248 standard; cDNA; 410 BP.  
XX  
XX  
AC ACH49248;  
XX  
XX  
DT 13-OCT-2003 (first entry)  
DE  
DE Human leukocyte cDNA #842.  
XX  
XX  
KW Human; ss; sequencing by hybridisation; SH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
OS Homo sapiens.  
XX  
XX  
FN US2003073623-A1.  
PD 17-APR-2003.  
XX  
XX  
PF 30-JUL-2001; 2001US-00918995.

XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI WPI; 2003-615964/58.  
XX  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX  
PS Claim 1; SEQ ID NO 36460; 44pp; English.  
XX  
XX  
CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH2789-ACH50831, whose sequence was  
CC determined by the technique of SH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030073623  
XX  
XX  
SQ Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 21; DB 8; Length 410;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CCTGTGCTCTATGCGTCTG 21  
DB 204 CCTGTGCTCTATGCGTCTG 184  
  
RESULT 10  
AAA43818/c  
ID AAA43818 standard; cDNA; 436 BP.  
XX  
XX  
AC AAA43818;  
XX  
XX  
DT 21-AUG-2000 (first entry)  
DE  
DE Mouse secreted expressed sequence tag SEQ ID NO:393.  
XX  
XX  
KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;  
KW antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
KW autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
KW infection; depression; psoriasis; ss.

```

OS Mus musculus.
XX WO200021991-A1.
XX PD 20-APR-2000.
XX PF 15-OCT-1999; 99WO-US024206.
XX PR 15-OCT-1998; 98US-0104436P.
XX PA (GENY ) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX PI Merberg D, Treacy M, Bowman MR;
XX WPI; 2000-317938/27.
XX Isolated polynucleotides, and encoded proteins, comprising secreted
XX expressed sequence tags (sESTs), useful for treating various disorders
XX such as autoimmune, infectious, and central nervous system disorders.
XX Claim 1; Page 306; 803pp; English.
XX AAA43426 to AAA45925 represent specifically claimed secreted expressed
XX sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue
XX sources. The sESTs can have a range of activities depending on the
XX tissues they were isolated from. The activities include: chemotactic;
XX proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;
XX haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;
XX antifungal; antiviral; antidiabetic; antiasthmatic; vulnery; antitumor;
XX osteoprotective; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;
XX cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be
XX used for gene therapy and in vaccines. The sESTs are useful as probes for
XX the identification and isolation of full-length cDNAs and genomic DNA
XX molecules which correspond to the sESTs. Proteins encoded by the sESTs
XX are useful in assays for determining biological activity and raising
XX antibodies. They may be useful for treatment of autoimmune disorders
XX (multiple sclerosis, insulin dependent diabetes), allergic conditions
XX (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
XX osteoporosis, osteoarthritis, central nervous system disorders
XX (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
XX disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
XX disease), tumours, bacterial, fungal or viral infections, depression and
XX psoriasis. AAA45926 to AAA45931 represent linker variants which are given
XX in the exemplification of the present invention
XX SQ Sequence 436 BP; 108 A; 113 C; 106 G; 109 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 3; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTGGTCTATGCGCTG 21
DB 226 CCTGTGGTCTATGCGCTG 206
RESULT 11
AAV86216/c
ID AAV86216 standard; cDNA; 466 BP.
XX AAV86216;
XX 27-APR-1999 (first entry)
XX EST clone O67.
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX tissue growth; activin; inhibitor; tumour invasion suppressor; EST; human;
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX

```

```

OS Homo sapiens.
XX WO9845435-A2.
XX PD 15-OCT-1998.
XX PF 10-APR-1998; 98WO-US006954.
XX PR 10-APR-1997; 97US-00835913.
XX PA (GENY ) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
XX PI Spaulding V, Agostino MJ;
XX WPI; 1999-070076/06.
XX New polynucleotides encoding human secreted proteins - derived from e.g.
XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,
XX pituitary, retina and colon cDNA libraries.
XX Claim 1; Page 160-161; 633pp; English.
XX This sequence represents an expressed sequence tag (EST), and is a
XX polynucleotide of the invention. The polynucleotides of the invention are
XX all secreted EST sequences isolated from a variety of human tissue
XX sources. The EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity, haemostatic
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX activity. The EST sequences are also stated to be useful for gene therapy
XX SQ Sequence 466 BP; 110 A; 122 C; 122 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTGGTCTATGCGCTG 21
DB 172 CCTGTGGTCTATGCGCTG 152
RESULT 12
AAC00117/c
ID AAC00117 standard; cDNA; 476 BP.
XX AAC00117;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 115.
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX

```

PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
DR P-PSDB; AAG00111.  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX Claim 1; SEQ ID NO 115; 71pp + Sequence Listing; English.  
PS  
XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors  
XX  
SQ Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;  
Query Match 100.0%; Score 21; DB 3; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTTGGTCTATCGCTCG 21  
DB 247 CCTGTTGGTCTATCGCTCG 227  
RESULT 13  
ACH49928/c 100.0%; Score 21; DB 3; Length 476;  
ID ACH49928 standard; cDNA; 490 BP.  
XX  
AC ACH49928;  
XX  
XX 13-OCT-2003 (first entry)  
DE Human leukocyte cDNA #1522.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX {DRMA/} DRMANAC R T.  
XX {LABA/} LABAT I.  
XX {STAC/} STACHE-CRAIN B.  
XX {DICK/} DICKSON M C.  
XX {JONE/} JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating

PT antisense DNA or RNA.  
XX  
PS Claim 1; SEQ ID NO 37140; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 490 BP; 112 A; 130 C; 125 G; 121 T; 0 U; 2 Other;  
Query Match 100.0%; Score 21; DB 8; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTTGGTCTATCGCTCG 21  
DB 277 CCTGTTGGTCTATCGCTCG 257  
RESULT 14  
AAC00116/c 100.0%; Score 21; DB 8; Length 490;  
ID AAC00116 standard; cDNA; 576 BP.  
XX  
AC AAC00116;  
XX  
XX 06-OCT-2000 (first entry)  
DE Human secreted protein 5' EST, SEQ ID NO: 114.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
XX EP1033401-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 21-FEB-2000; 2000EP-00200610.  
XX  
XX 26-FEB-1999; 99US-0122487P.  
XX  
XX {GEST } GENSET.  
XX  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
XX WPI; 2000-500381/45.  
XX P-PSDB; AAG00110.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX Claim 1; SEQ ID NO 114; 71pp + Sequence Listing; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors  
XX  
SQ Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;  
Query Match 100.0%; Score 21; DB 3; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTTGGTCTATCGCTCG 21  
DB 247 CCTGTTGGTCTATCGCTCG 227



CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors  
 CC  
 CC SQ Sequence 576 BP; 131 A; 158 C; 141 G; 145 T; 0 U; 1 Other;

Query Match 100.0%; Score 21; DB 3; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATGCGTCTG 21  
 |||||  
 Db 347 CCTGTGCTCTATGCGTCTG 327

RESULT 15  
 AAQ25060/c  
 ID AAQ25060 standard; DNA; 1202 BP.

XX AC AAQ25060;

XX DT 15-NOV-1992 (first entry)

XX DE PSBAlpha-318 clone.

XX KW RFLP; probe p29G8; beta-lymphoblastoid cell line; IG2; T5-1; 6.3.6;

XX KW transplant; transfection; paternity; ss.

XX OS Synthetic.

XX PN US5110920-A.

XX PD 05-MAY-1992.

XX PF 05-DEC-1984; 84US-00678255.

XX PR 22-JAN-1982; 82US-00341902.

XX PR 07-JAN-1983; 83US-00456373.

XX PR 30-AUG-1988; 88US-00238619.

XX PA (CETU ) CETUS CORP.

XX PI Brlich HA;

XX DR WPI; 1992-175244/21.

XX PT New DNA probes specific to single class II HLA locus - useful in HLA

XX PT typing e.g. to evaluate paternity and transplant or transfection

XX PT compatibility and to diagnose disease susceptibility.

XX PS Disclosure; Page 11; 21pp; English.

XX CC The sequence given is a pSAlpha-318 clone which was derived from a beta-  
 CC lymphoblastoid cell line IG2 cDNA library using a probe designated p29G8.  
 CC This probe bound to sequences distinct from those which lead to the  
 CC elucidation of HLA-B\*34 (see also AAQ25059). A genomic blot pattern with  
 CC DNA from the cell lines T5-1 and its HLA hemizygous derivative 6.3.6  
 CC indicates that the p29G8 locus maps within the HLA region. p29G8 has been  
 CC found to be a HLA-SAlpha clone and could be used to isolate the given  
 CC sequence. p29G8 could be useful in HLA typing based on RFLPs. It can be  
 CC utilized in paternity disputes or for determining transplant or transfection  
 CC compatibility. It can also be used to make disease correlations to  
 CC diagnose diseases or predict susceptibility to diseases

XX SQ Sequence 1202 BP; 289 A; 336 C; 277 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 1202;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATGCGTCTG 21  
 |||||  
 Db 232 CCTGTGCTCTATGCGTCTG 212

RESULT 16

AA31123/c

ID AAS31123 standard; cDNA; 1259 BP.

XX AC AAS31123;

XX DT 04-DEC-2001 (first entry)

XX DE Human diagnostic and therapeutic polynucleotide (DITHP) #138.

XX KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;

XX KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;

XX KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;

XX KW respiratory disorder; ss.

XX OS Homo sapiens.

XX PN WO200162927-A2.

XX PD 30-AUG-2001.

XX PF 21-FEB-2001; 2001WO-US006059.

XX PR 24-FEB-2000; 2000US-0184693P.

XX PR 24-FEB-2000; 2000US-0184697P.

XX PR 24-FEB-2000; 2000US-0184698P.

XX PR 24-FEB-2000; 2000US-0184768P.

XX PR 24-FEB-2000; 2000US-0184769P.

XX PR 24-FEB-2000; 2000US-0184770P.

XX PR 24-FEB-2000; 2000US-0184771P.

XX PR 24-FEB-2000; 2000US-0184772P.

XX PR 24-FEB-2000; 2000US-0184773P.

XX PR 24-FEB-2000; 2000US-0184774P.

XX PR 24-FEB-2000; 2000US-0184776P.

XX PR 24-FEB-2000; 2000US-0184777P.

XX PR 24-FEB-2000; 2000US-0184797P.

XX PR 24-FEB-2000; 2000US-0184813P.

XX PR 24-FEB-2000; 2000US-0184837P.

XX PR 24-FEB-2000; 2000US-0184841P.

XX PR 24-FEB-2000; 2000US-0185213P.

XX PR 24-FEB-2000; 2000US-0185216P.

XX PR 12-MAY-2000; 2000US-0203785P.

XX PR 15-MAY-2000; 2000US-0204226P.

XX PR 16-MAY-2000; 2000US-0204525P.

XX PR 16-MAY-2000; 2000US-0204821P.

XX PR 16-MAY-2000; 2000US-0204908P.

XX PR 16-MAY-2000; 2000US-0205232P.

XX PR 17-MAY-2000; 2000US-0204815P.

XX PR 17-MAY-2000; 2000US-0204863P.

XX PR 17-MAY-2000; 2000US-0205221P.

XX PR 17-MAY-2000; 2000US-0205285P.

XX PR 17-MAY-2000; 2000US-0205286P.

XX PR 17-MAY-2000; 2000US-0205287P.

XX PR 17-MAY-2000; 2000US-0205323P.

XX PR 17-MAY-2000; 2000US-0205324P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;

XX PI Chen A, D'sa SA, Ameshey S, Dahl CR, Dam TC, Daniels SE, Dufour GE;

XX PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF;

XX PI Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;

XX PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;

XX PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;

XX WPI; 2001-502867/55.  
 DR P-PSDB; AAU19552.  
 XX  
 PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 XX  
 XX Claim 1; Page 361; 522pp; English.  
 XX  
 XX The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DITHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide  
 CC expression by rectifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DITHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DITHPs and in assays to identify modulators of DITHP  
 CC expression and activity. The anti-DITHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DITHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay  
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic  
 CC (DITHP) polynucleotides of the invention  
 XX  
 XX Sequence 1259 BP; 293 A; 350 C; 300 G; 316 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 21; DB 4; Length 1259;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATGCGTCTG 21  
 |||||  
 Db 367 CCCTGTTGGTCTATGCGTCTG 347  
 |||||

RESULT 17  
 AAF18332/c  
 ID AAF18332 standard; DNA; 1348 BP.  
 AC AAF18332;  
 XX  
 XX 14-MAR-2001 (first entry)  
 DT  
 XX Lung cancer associated polynucleotide sequence SEQ ID 351.  
 DE  
 XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WC200055180-A2.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 08-MAR-2000; 2000WO-US005918.  
 PF  
 XX 12-MAR-1999; 99US-0124270P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX

PA (ROSE/) ROSEN C A.  
 XX  
 XX Ruben SM;  
 XX  
 DR WPI; 2000-587514/55.  
 DR P-PSDB; AAB58456.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer.  
 XX  
 XX Claim 1; Page 808-809; 1425pp; English.  
 PS  
 XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; of antibacterial  
 CC activity. The invention also includes antibodies specific for the protein  
 CC or polynucleotide sequences. The lung cancer associated polynucleotide  
 CC sequences may be used for detection of lung cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The proteins may be used to treat disorders such as  
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,  
 CC cardiovascular, renal, and proliferative disorders. The proteins may also  
 CC be used in the treatment of wounds and infectious diseases.  
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are  
 CC used in the course of the invention for the identification and  
 CC characterisation of the polynucleotide and protein sequences  
 XX  
 XX Sequence 1348 BP; 335 A; 373 C; 307 G; 324 T; 0 U; 9 Other;  
 SQ  
 Query Match 100.0%; Score 21; DB 3; Length 1348;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATGCGTCTG 21  
 |||||  
 Db 305 CCCTGTTGGTCTATGCGTCTG 285  
 |||||

RESULT 18  
 AABK64796/c  
 ID AABK64796 standard; DNA; 14646 BP.  
 XX  
 XX AABK64796;  
 AC  
 XX 18-JUN-2002 (first entry)  
 DT  
 XX Human benign prostatic hyperplasia gene #691.  
 DE  
 XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WC200212440-A2.  
 PN  
 XX 14-FEB-2002.  
 PD  
 XX 07-AUG-2001; 2001WO-US024708.  
 PF  
 XX 07-AUG-2000; 2000US-0223323P.  
 PR  
 XX 05-JUN-2001; 2001US-00873319.  
 XX  
 XX (GENE-) GENE LOGIC INC.  
 PA  
 XX (NISR) JAPAN TOBACCO INC.  
 XX  
 XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
 PI  
 XX WPI; 2002-257476/30.  
 DR  
 XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 PT

PT detecting expression levels of one or more genes in prostate cells from  
 PT patient that are differentially regulated compared to normal prostate  
 PT cells.

PS Disclosure; Page 373-377; 444pp; English.

XX The invention relates to a method of diagnosing (I) the onset or  
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
 CC or identifying an agent that modulates the onset or progression of BPH.  
 CC The method is based on changes in gene expression in BPH tissue isolated  
 CC from patients exhibiting different clinical states of prostate  
 CC hyperplasia as compared to normal prostate tissue. (I) comprises  
 CC detecting the expression levels of one or more genes in prostate cells  
 CC from the subject that are differentially regulated compared to normal  
 CC prostate cells. (II) comprises preparing a first gene expression profile  
 CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles. (I)  
 CC is useful for diagnosing the onset or progression of BPH. (II) is useful  
 CC for identifying an agent that modulates the onset or progression of BPH.  
 CC The methods are useful to present information identifying the expression  
 CC level in a tissue or cells, by comparing the expression level of genes  
 CC given in the specification in the tissue or cells to the level of  
 CC expression of gene in the database, and displaying the expression levels  
 CC of at least one gene in the tissue or cell sample compared to the  
 CC expression level in BPH. Agents using (II) are useful for treating BPH or  
 CC prostate cancer. ABX64106-ABX64860 represent human benign prostatic  
 CC hyperplasia gene sequences of the invention

XX Sequence 14646 BP; 4552 A; 2957 C; 2741 G; 4395 T; 0 U; 1 Other;

Query Match 100.0%; Score 21; DB 6; Length 14646;  
 Best Local Similarity 100.0%; Pred.No. 2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTTGGTCTATCGGCTG 21

DB 6386 CCTGTTGGTCTATCGGCTG 6366

RESULT 19

AA48203/C

ID AAL48203 standard; DNA; 20 BP.

XX AC AAL48203;

XX AC AAL48203;

XX 01-OCT-2002 (first entry)

XX Human HLA DPAl locus polymorphism primer sequence #3.

XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;

XX Flow cytometry; human; DPAl; DPB1; PCR; primer; ss.

XX Homo sapiens.

XX WO200194639-A1.

XX 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US018590.

XX 08-JUN-2000; 2000US-0210759P.

XX (REGC ) UNIV CALIFORNIA.

XX White PS, Torney DC;

XX WPI; 2002-566450/60.

XX Identifying sequences useful as address/capture tags for flow cytometry

XX based minisequencing, by generating tag sequences and rejecting

XX PT based on certain parameters e.g. sequences which form stable hairpins.

XX

PS Disclosure; Page 14; 35pp; English.

XX The present invention relates to a method of identifying sequences useful  
 CC as address/capture tags, involving rejecting sequences having common sub-  
 CC sequences with a sub-sequence length greater than specified number of  
 CC bases, and sequences which can form stable hairpins and stable dimers  
 CC from a sample of oligonucleotides, and selecting those sequences in the  
 CC sample that would hybridize to their respective complements with a high  
 CC degree of specificity. The method is useful for identifying a set of  
 CC sequences useful as address/capture tags which can be used for  
 CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
 CC cytometry assay. The present sequence is a primer tag described in the  
 CC exemplification of the invention

XX Sequence 20 BP; 8 A; 6 C; 5 G; 1 T; 0 U; 0 Other;

Query Match 95.2%; Score 20; DB 6; Length 20;

Best Local Similarity 100.0%; Pred.No. 3.2;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTTGGTCTATCGGCTG 21

DB 20 CCTGTTGGTCTATCGGCTG 1

RESULT 20

AAV86130/C

ID AAV86130 standard; cDNA; 396 BP.

XX AC AAV86130;

XX DT 27-APR-1999 (first entry)

XX DE EST clone H45.

XX Expressed sequence tag; secreted protein; haematopoiesis regulator;

XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;

XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;

XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX Homo sapiens.

XX WO9845435-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98WO-US006954.

XX 10-APR-1997; 97US-00835913.

XX (GENY ) GENETICS INST INC.

XX Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

XX Spaulding V, Agostino MJ;

XX WPI; 1999-070076/06.

XX New polynucleotides encoding human secreted proteins - derived from e.g.

XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,

XX pituitary, retina and colon cDNA libraries.

XX Claim 1; Page 131; 633pp; English.

XX This sequence represents an expressed sequence tag (EST), and is a

XX polynucleotide of the invention. The polynucleotides of the invention are

XX all secreted EST sequences isolated from a variety of human tissue

XX sources. The EST sequences and proteins encoded by them are predicted to

XX have useful biological activities which would make them suitable for

XX treating, preventing or ameliorating medical conditions in humans and

XX animals, although no supporting data is given. Suggested activities

XX include nutritional activity, immune stimulating or suppressing activity,

XX haematopoiesis regulating activity, tissue growth activity,

XX chemotactic/chemokinetic activity, haemostatic

XX activin/inhibin activity,

CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy

XX  
 SQ Sequence 396 BP; 100 A; 97 C; 101 G; 98 T; 0 U; 0 Other;  
 Query Match 92.4%; Score 19.4; DB 2; Length 396;  
 Best Local Similarity 95.2%; Pred. No. 8.4;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTCTGTGCTCATGCGTCTG 21  
 Db 204 CCTCTGTGCTCATGCGTCTG 184  
 |||||

RESULT 21  
 AAZ90198  
 ID AAZ90198 standard; DNA; 4233 BP.

AC AAZ90198;

XX 19-MAY-2000 (first entry)

DE Rat mdrlb2 (multispecific drug transporter) nucleotide sequence.

XX Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;  
 KW cardiovascular disease; central nervous system disorder;  
 KW autoimmune disease; kidney disease; drug formulation; ss.

XX Rattus sp.

XX US6025160-A.

PN 15-FEB-2000.

XX 22-JUL-1998; 98US-00120513.

XX 22-JUL-1998; 98US-00120513.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;

XX WPI; 2000-181810/16.

XX P-PSDB; AA78879.

XX Isolated polynucleotide encoding a rat mdrlb2 drug transporter  
 PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,  
 PT central nervous system disorders.

XX Claim 7; Col 13-17; 17pp; English.

XX This sequence represents a polynucleotide encoding a 1275 amino acid rat  
 CC mdrlb2 multispecific drug transporter polypeptide. The mdrlb2 nucleotide  
 CC sequence has cytostatic, antiinflammatory, cardiant, neuroprotective,  
 CC immunosuppressive and nephrotropic activity. Understanding the  
 CC functioning of the mdrlb2 polynucleotide and protein in transgenic animal  
 CC models is useful for treating and preventing diseases such as cancer,  
 CC inflammation, cardiovascular disease, central nervous system disorders,  
 CC autoimmune disease, and kidney disease. The use of the protein in cell  
 CC based, membrane based, or binding assays may enhance drug formulation,  
 CC selection of formulation excipients and compound design

XX Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 U; 0 Other;

XX Query Match 90.5%; Score 19; DB 3; Length 4233;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTGCTCATGCGTCT 20

Db 928 CCTGTGCTCATGCGTCT 946  
 |||||

RESULT 22  
 AAF27498  
 ID AAF27498 standard; cDNA; 4233 BP.

XX AAF27498;

AC 25-APR-2001 (first entry)

DE Rat mdrlb2 multidrug resistance cDNA, SEQ ID NO:1.

XX Rat; multidrug resistance; mdrlb2; multi-drug transporter family;  
 KW drug screening; pharmacokinetic analysis; oral absorption;  
 KW formulation design; bioavailability; transgenic animal; knockout animal;  
 KW inflammation; cardiovascular disease; central nervous system disorder;  
 KW cancer; autoimmune disorder; kidney disease; ss.

XX Rattus sp.

OS US6169166-B1.

PN 02-JAN-2001.

XX 29-NOV-1999; 99US-00450105.

XX 22-JUL-1998; 98US-00120513.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX (SMIK ) SMITHKLINE BEECHAM PLC.

XX Chenery RJ, Ellens H, Brun KA, Yue L, Feild JA;

XX WPI; 2001-158280/16.

XX P-PSDB; AAB60409.

XX New rat multidrug resistance protein (mdrlb2), useful for screening or  
 PT identifying compounds that are (anti)agonists/inhibitors of the ratmdrlb2,  
 PT as well as compounds with optimal development characteristics.

XX Disclosure; Col 13-14; 13pp; English.

XX The invention relates to a novel rat multidrug resistance (mdr) protein,  
 CC mdrlb2 (AAB60409). Multi-specific drug transporter family proteins are  
 CC present in cell which have a barrier function, such as intestinal  
 CC epithelial cells, brain microvessel endothelial cells, kidney epithelial  
 CC cells, and liver hepatocytes, and are also expressed by certain cancer  
 CC cells. The rat mdrlb2 protein is useful for screening or identifying  
 CC compounds that are agonists or antagonists of mdrlb2 activity. It may  
 CC also be used to establish assays to predict oral absorption and  
 CC pharmacokinetics of drugs in humans, and thus enhance the design of  
 CC formulations through the identification of compounds with optimal  
 CC development characteristics (i.e., high oral bioavailability, and reduced  
 CC food effects), specifically to avoid interactions with human mdr-1.  
 CC Transgenic and knockout animals created using DNA encoding the rat mdrlb2  
 CC may be used to gain an insight into treating and preventing human  
 CC diseases such as cancer, inflammation, cardiovascular disease, central  
 CC nervous system disorders, autoimmune disorders and kidney disease. The  
 CC present sequence represents cDNA encoding rat mdrlb2

XX Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 U; 0 Other;

XX Query Match 90.5%; Score 19; DB 4; Length 4233;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTGCTCATGCGTCT 20

Db 928 CCTGTGCTCATGCGTCT 946  
 |||||

RESULT 23

ABR63517

ID ABK63517 standard; cDNA; 4254 BP.  
XX  
AC ABK63517;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Rat sequence differentially expressed in response to a hepatotoxin #1424.  
XX  
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
XX differential expression; centrilobular necrosis; steatosis.  
XX  
OS Rattus norvegicus.  
XX  
FN WO200210453-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 30-JUL-2001; 2001WO-US023872.  
XX  
PR 31-JUL-2000; 2000US-0222040P.  
PR 02-NOV-2000; 2000US-0244880P.  
PR 11-MAY-2001; 2001US-0290029P.  
PR 15-MAY-2001; 2001US-0290645P.  
PR 22-MAY-2001; 2001US-0292336P.  
PR 06-JUN-2001; 2001US-0295798P.  
PR 13-JUN-2001; 2001US-0297457P.  
PR 19-JUN-2001; 2001US-0298884P.  
PR 09-JUL-2001; 2001US-0303459P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
XX  
DR WPI; 2002-241625/29.  
XX  
PT Predicting toxic effects of compounds or the progression of these toxic  
XX effects by determining the changes in gene expression in tissues or cells  
XX exposed to the toxin and comparing these to gene expression in unexposed  
XX tissues or cells.  
XX  
PS Claim 1; SEQ ID NO 1424; 239pp; English.  
XX  
CC The invention relates to methods for predicting toxic effects of  
XX compounds or the progression of these toxic effects by determining the  
XX global changes in gene expression in tissues or cells exposed to the  
XX toxin and comparing these to gene expression in unexposed tissues or  
XX cells. Also included are methods of predicting at least one toxic effect  
XX of a compound or progression of a toxic effect, preferably the  
XX hepatotoxicity of a compound, comprising detecting the level of  
XX expression in a tissue or cell sample exposed to the compound of two or  
XX more genes listed in the specification, where differential expression of  
XX the genes is indicative of at least one toxic effect or progression. The  
XX method can also be used to identify an agent which modulates the toxic  
XX response and predict cellular pathways that a compound modulates in a  
XX cell. The methods utilise a set of at least two probes (on a solid  
XX support in kit form), where each of the probes comprises a sequence that  
XX specifically hybridises to a gene listed in the specification, a computer  
XX system comprising a database containing information identifying the  
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a  
XX set of genes comprising at least two genes listed in the specification,  
XX and a user interface to view the information used to present information  
XX identifying the expression level in a tissue or cell of at least one gene  
XX listed in the specification. The method is useful for elucidating global  
XX changes in gene expression and for identifying toxicity markers in  
XX tissues or cell exposed to a known toxin. The genes may be used as  
XX toxicity markers in drug screening and toxicity assays. The genes and  
XX gene expression information may be used as diagnostic markers for the  
XX prediction or identification of the physiological state of tissue or cell  
XX sample that has been exposed to a compound or agent. Hepatotoxicity is  
XX characterised by centrilobular necrosis and steatosis. The present  
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
XX which is differentially expressed in response to a hepatotoxic agent

SQ Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 U; 0 Other;  
Query Match 90.5%; Score 19; DB 6; Length 4254;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 2 CCTGTGGTCTATGCGTCT 20  
Db 1008 CCTGTGGTCTATGCGTCT 1026  
RESULT 24  
ABT41782  
ID ABT41782 standard; DNA; 4254 BP.  
XX  
AC ABT41782;  
XX  
DT 26-JUN-2003 (first entry)  
XX  
DE Toxicity modelling related rat gene SEQ ID No 1484.  
XX  
KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;  
XX database; drug screening; toxicity assay; rat; ds.  
XX  
OS Rattus norvegicus.  
XX  
FN WO200295000-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 22-MAY-2002; 2002WO-US016173.  
XX  
PR 22-MAY-2001; 2001US-0292335P.  
PR 13-JUN-2001; 2001US-0297523P.  
PR 19-JUN-2001; 2001US-0298925P.  
PR 10-JUL-2001; 2001US-0303807P.  
PR 10-JUL-2001; 2001US-0303808P.  
PR 10-JUL-2001; 2001US-0303810P.  
PR 28-AUG-2001; 2001US-0315047P.  
PR 27-SEP-2001; 2001US-0324928P.  
PR 22-OCT-2001; 2001US-0330462P.  
PR 01-NOV-2001; 2001US-0330867P.  
PR 21-NOV-2001; 2001US-0331805P.  
PR 06-DEC-2001; 2001US-0336144P.  
PR 19-DEC-2001; 2001US-0340873P.  
PR 21-FEB-2002; 2002US-0357842P.  
PR 21-FEB-2002; 2002US-0357843P.  
PR 15-MAR-2002; 2002US-0364134P.  
PR 08-APR-2002; 2002US-0370144P.  
PR 08-APR-2002; 2002US-0370206P.  
PR 08-APR-2002; 2002US-0370247P.  
PR 17-APR-2002; 2002US-0372794P.  
PR 21-APR-2002; 2002US-0371679P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;  
XX  
DR WPI; 2003-148464/14.  
XX  
PT Predicting at least one toxic effect of a compound, useful for toxicity  
XX modelling, comprises preparing a gene expression profile of a tissue or  
XX cell sample exposed to the compound, and comparing the gene expression  
XX profile to a database.  
XX  
PS Example 4; Page; 446pp; English.  
XX  
CC The invention relates to a novel method of predicting at least one toxic  
XX effect of a compound. The method comprises a gene expression profile of a  
XX tissue or cell sample exposed to the compound, and comparing the gene  
XX expression profile to a database comprising at least part of the data or  
XX information given in the specification. The methods are useful for

CC predicting at least one toxic effect of a compound, predicting the  
CC progression of a toxic effect of a compound, predicting the renal  
CC toxicity of a compound, or identifying toxicity markers in tissues or  
CC cells exposed to known renal toxin. The genes are useful as toxicity  
CC markers in drug screening and toxicity assays, in monitoring disease or  
CC physiological states, or disease progression. This polynucleotide  
CC represents a rat DNA sequence relating to the toxic effect database  
CC described in the specification. NOTE: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the World Intellectual Property  
CC Organization

XX Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 U; 0 Other;

SQ Query Match 90.5%; Score 19; DB 7; Length 4254;

Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTGGTCTATGCGTCT 20

Db 1008 CCTGTGGTCTATGCGTCT 1026

RESULT 25

AAI99682\_03

Continuation (4 of 45) of AAI99682 from base 300001 (Mycobacterium tuberculosis strain H

WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682

Fragment Name Begin End

WP AAI99682\_00 1 110000

WP AAI99682\_01 100001 210000

WP AAI99682\_02 200001 310000

WP AAI99682\_03 300001 410000

WP AAI99682\_04 400001 510000

WP AAI99682\_05 500001 610000

WP AAI99682\_06 600001 710000

WP AAI99682\_07 700001 810000

WP AAI99682\_08 800001 910000

WP AAI99682\_09 900001 1010000

WP AAI99682\_10 1000001 1110000

WP AAI99682\_11 1100001 1210000

WP AAI99682\_12 1200001 1310000

WP AAI99682\_13 1300001 1410000

WP AAI99682\_14 1400001 1510000

WP AAI99682\_15 1500001 1610000

WP AAI99682\_16 1600001 1710000

WP AAI99682\_17 1700001 1810000

WP AAI99682\_18 1800001 1910000

WP AAI99682\_19 1900001 2010000

WP AAI99682\_20 2000001 2110000

WP AAI99682\_21 2100001 2210000

WP AAI99682\_22 2200001 2310000

WP AAI99682\_23 2300001 2410000

WP AAI99682\_24 2400001 2510000

WP AAI99682\_25 2500001 2610000

WP AAI99682\_26 2600001 2710000

WP AAI99682\_27 2700001 2810000

WP AAI99682\_28 2800001 2910000

WP AAI99682\_29 2900001 3010000

WP AAI99682\_30 3000001 3110000

WP AAI99682\_31 3100001 3210000

WP AAI99682\_32 3200001 3310000

WP AAI99682\_33 3300001 3410000

WP AAI99682\_34 3400001 3510000

WP AAI99682\_35 3500001 3610000

WP AAI99682\_36 3600001 3710000

WP AAI99682\_37 3700001 3810000

WP AAI99682\_38 3800001 3910000

WP AAI99682\_39 3900001 4010000

WP AAI99682\_40 4000001 4110000

WP AAI99682\_41 4100001 4210000

WP AAI99682\_42 4200001 4310000

WP AAI99682\_43 4300001 4410000

WP AAI99682\_44 4400001 4411529

Query Match 84.8%; Score 17.8; DB 4; Length 110000;  
Best Local Similarity 90.5%; Pred. No. 90;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21

Db 81808 CCTGTGGTCTATGCGTCTG 81828

RESULT 26

AAI99683\_03

Continuation (4 of 44) of AAI99683 from base 300001 (Mycobacterium tuberculosis strain H

WP Sequence split into 44 fragments LOCUS AAI99683 Accession Aai99683

Fragment Name Begin End

WP AAI99683\_00 1 110000

WP AAI99683\_01 100001 210000

WP AAI99683\_02 200001 310000

WP AAI99683\_03 300001 410000

WP AAI99683\_04 400001 510000

WP AAI99683\_05 500001 610000

WP AAI99683\_06 600001 710000

WP AAI99683\_07 700001 810000

WP AAI99683\_08 800001 910000

WP AAI99683\_09 900001 1010000

WP AAI99683\_10 1000001 1110000

WP AAI99683\_11 1100001 1210000

WP AAI99683\_12 1200001 1310000

WP AAI99683\_13 1300001 1410000

WP AAI99683\_14 1400001 1510000

WP AAI99683\_15 1500001 1610000

WP AAI99683\_16 1600001 1710000

WP AAI99683\_17 1700001 1810000

WP AAI99683\_18 1800001 1910000

WP AAI99683\_19 1900001 2010000

WP AAI99683\_20 2000001 2110000

WP AAI99683\_21 2100001 2210000

WP AAI99683\_22 2200001 2310000

WP AAI99683\_23 2300001 2410000

WP AAI99683\_24 2400001 2510000

WP AAI99683\_25 2500001 2610000

WP AAI99683\_26 2600001 2710000

WP AAI99683\_27 2700001 2810000

WP AAI99683\_28 2800001 2910000

WP AAI99683\_29 2900001 3010000

WP AAI99683\_30 3000001 3110000

WP AAI99683\_31 3100001 3210000

WP AAI99683\_32 3200001 3310000

WP AAI99683\_33 3300001 3410000

WP AAI99683\_34 3400001 3510000

WP AAI99683\_35 3500001 3610000

WP AAI99683\_36 3600001 3710000

WP AAI99683\_37 3700001 3810000

WP AAI99683\_38 3800001 3910000

WP AAI99683\_39 3900001 4010000

WP AAI99683\_40 4000001 4110000

WP AAI99683\_41 4100001 4210000

WP AAI99683\_42 4200001 4310000

WP AAI99683\_43 4300001 4403765

Query Match 84.8%; Score 17.8; DB 4; Length 110000;  
Best Local Similarity 90.5%; Pred. No. 90;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21

Db 81864 CCTGTGGTCTATGCGTCTG 81884

RESULT 27

ADA02564

ID ADA02564 standard; DNA; 96596 BP.

XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 : Search time 28.5353 Seconds  
(without alignments)

408.405 Million cell updates/sec

Title: US-09-877-819B-40

Perfect score: 21

Sequence: 1 cccgtgtgctatgctgctg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	90.5	4233	3	US-09-120-513-1
2	19	90.5	4233	3	US-09-450-105-1
3	17.8	84.8	1860	4	US-09-894-844-65
4	17.8	84.8	4403765	3	US-09-103-840A-2
5	17.8	84.8	4411529	3	US-09-103-840A-1
6	16.2	77.1	7785	2	US-08-276-967-1
7	16	76.2	148567	4	US-09-801-876B-3
8	16	76.2	148567	4	US-10-254-869-3
9	15.8	75.2	112132	4	US-09-741-150-3
10	15.8	75.2	112132	4	US-10-160-187-3
11	15.4	73.3	1068	4	US-09-328-352-2430
12	15.2	72.4	1178	3	US-08-861-774E-35
13	14.8	70.5	483	4	US-09-489-039A-5676
14	14.8	70.5	1332	4	US-09-711-164-143
15	14.8	70.5	1335	4	US-09-489-039A-3751
16	14.8	70.5	3088	1	US-08-418-44A-1
17	14.8	70.5	20986	4	US-08-961-527-54
18	14.8	70.5	36138	4	US-08-311-731A-136
19	14.6	69.5	271	4	US-09-313-294A-33
20	14.6	69.5	442	4	US-09-280-116-164
21	14.6	69.5	525	4	US-09-621-976-17799
22	14.6	69.5	576	4	US-09-252-991A-10362
23	14.6	69.5	590	1	US-08-466-033-116
24	14.6	69.5	590	1	US-08-444-733-116
25	14.6	69.5	590	2	US-08-464-134-116
26	14.6	69.5	590	2	US-08-461-361-116
27	14.6	69.5	590	2	US-08-485-910-116

3	US-09-328-111-576	596	69.5	14.6	Sequence 576, App
3	US-09-386-493-4	944	69.5	14.6	Sequence 4, Appli
3	US-09-540-236-1082	1026	69.5	14.6	Sequence 1082, Ap
4	US-08-943-607-17	1218	69.5	14.6	Sequence 17, Appli
3	US-08-932-823A-1	1302	69.5	14.6	Sequence 1, Appli
4	US-09-489-039A-5151	1446	69.5	14.6	Sequence 5151, Ap
4	US-09-540-224-3	1472	69.5	14.6	Sequence 3, Appli
4	US-09-564-596D-52	1472	69.5	14.6	Sequence 52, Appli
4	US-09-252-991A-10612	1497	69.5	14.6	Sequence 3, Appli
4	US-09-808-977A-3	1638	69.5	14.6	Sequence 10612, A
4	US-09-489-039A-3879	2268	69.5	14.6	Sequence 3879, Ap
4	US-09-620-312D-909	2731	69.5	14.6	Sequence 909, App
4	US-09-016-434-1344	4665	69.5	14.6	Sequence 1344, Ap
4	US-09-252-991A-10243	7724	69.5	14.6	Sequence 10243, A
1	US-08-466-049-1	9103	69.5	14.6	Sequence 1, Appli
1	US-08-466-033-182	9103	69.5	14.6	Sequence 182, App
2	US-08-444-733-182	9103	69.5	14.6	Sequence 182, App
2	US-08-464-134-182	9103	69.5	14.6	Sequence 182, App
2	US-08-461-361-182	9103	69.5	14.6	Sequence 182, App
2	US-08-485-910-182	9103	69.5	14.6	Sequence 182, App
5	PCT-US95-06266-156	9327	69.5	14.6	Sequence 156, App
1	US-08-466-033-234	9327	69.5	14.6	Sequence 234, App
1	US-08-444-733-234	9327	69.5	14.6	Sequence 234, App
2	US-08-464-134-234	9327	69.5	14.6	Sequence 234, App
2	US-08-461-361-234	9327	69.5	14.6	Sequence 234, App
2	US-08-485-910-234	9327	69.5	14.6	Sequence 234, App
1	US-08-638-911A-1	9391	69.5	14.6	Sequence 1, Appli
5	PCT-US95-06266-14	9391	69.5	14.6	Sequence 14, Appli
1	US-08-466-033-14	9392	69.5	14.6	Sequence 14, Appli
1	US-08-444-733-14	9392	69.5	14.6	Sequence 14, Appli
2	US-08-464-134-14	9392	69.5	14.6	Sequence 14, Appli
2	US-08-461-361-14	9392	69.5	14.6	Sequence 14, Appli
2	US-08-485-910-14	9392	69.5	14.6	Sequence 14, Appli
3	US-09-096-942-1	16075	69.5	14.6	Sequence 1, Appli
3	US-09-096-867-1	16075	69.5	14.6	Sequence 1, Appli
4	US-09-596-002-10	19988	69.5	14.6	Sequence 10, Appli
4	US-09-716-865-23	55216	69.5	14.6	Sequence 23, Appli
4	US-09-107-532A-1266	768	69.5	14.4	Sequence 1266, Ap
4	US-09-107-532A-1317	894	69.5	14.4	Sequence 1317, Ap
3	US-09-018-584A-37	1200	69.5	14.4	Sequence 37, Appli
4	US-09-252-991A-2121	1233	69.5	14.4	Sequence 2121, Ap
4	US-09-252-991A-1954	1395	69.5	14.4	Sequence 1954, Ap
4	US-09-539-333D-167	3001	69.5	14.4	Sequence 167, App
4	US-09-023-655-664	5304	69.5	14.4	Sequence 664, App
4	US-09-820-001-3	22067	69.5	14.4	Sequence 3, Appli
4	US-09-146-053-3	50000	69.5	14.4	Sequence 3, Appli
3	US-03-078-294-4	80246	69.5	14.4	Sequence 4, Appli
3	US-03-497-855A-38	80595	69.5	14.4	Sequence 38, Appli
4	US-09-734-674-3	128779	69.5	14.4	Sequence 3, Appli
4	US-09-751-389-3	786431	69.5	14.4	Sequence 3, Appli
4	US-09-252-991A-2817	465	69.5	14.2	Sequence 2817, Ap
4	US-09-134-000C-1037	951	69.5	14.2	Sequence 1037, Ap
4	US-09-671-317-242	1001	69.5	14.2	Sequence 241, App
4	US-09-671-317-242	1001	69.5	14.2	Sequence 242, App
4	US-09-671-317-243	1001	69.5	14.2	Sequence 243, App
4	US-09-671-317-243	1001	69.5	14.2	Sequence 244, App
4	US-09-671-317-245	1001	69.5	14.2	Sequence 245, App
2	US-09-014-969-6	1045	69.5	14.2	Sequence 6, Appli
3	US-03-347-798-7	1713	69.5	14.2	Sequence 7, Appli
4	US-09-252-991A-3061	3621	69.5	14.2	Sequence 3061, Ap
4	US-09-107-532A-1587	3729	69.5	14.2	Sequence 1587, Ap
4	US-09-976-594-792	3902	69.5	14.2	Sequence 792, App
4	US-09-762-195-1	3988	69.5	14.2	Sequence 1, Appli
4	US-09-367-750-1	4066	69.5	14.2	Sequence 1, Appli
2	US-08-784-649A-1	4264	69.5	14.2	Sequence 1, Appli
2	US-08-784-649A-5	4264	69.5	14.2	Sequence 5, Appli
4	US-09-484-970B-4	4359	69.5	14.2	Sequence 2, Appli
1	US-08-181-471-2	4646	69.5	14.2	Sequence 1167, Ap
4	US-09-023-655-1167	4646	69.5	14.2	Sequence 18, Appli
2	US-08-583-276-18	4669	69.5	14.2	Sequence 1, Appli
4	US-08-752-447-1	4669	69.5	14.2	Sequence 1, Appli
4	US-09-316-167-1	4669	69.5	14.2	Sequence 1, Appli

101 14.2 67.6 4669 4 US-09-397-233-1 Sequence 1, Appli  
102 14.2 67.6 4669 6 5206352-3 Patent No. 5206352  
103 14.2 67.6 5125 1 US-08-094-948A-4 Sequence 4, Appli  
104 14.2 67.6 5125 5 PCT-US96-09319-4 Sequence 4, Appli  
105 14.2 67.6 5125 2 US-08-793-610-5 Sequence 5, Appli  
106 14.2 67.6 8630 4 US-09-306-417-2 Sequence 1, Appli  
107 14.2 67.6 8630 4 US-09-306-417-2 Sequence 2, Appli  
108 14.2 67.6 9318 2 US-08-793-610-6 Sequence 6, Appli  
109 14.2 67.6 12734 4 US-09-344-456-1 Sequence 1, Appli  
110 14.2 67.6 14078 3 US-09-433-262-1 Sequence 1, Appli  
111 14.2 67.6 14078 4 US-09-702-330-1 Sequence 1, Appli  
112 14.2 67.6 14561 4 US-09-392-714-1 Sequence 1, Appli  
113 14.2 67.6 14578 3 US-08-859-694-1 Sequence 1, Appli  
114 13.8 65.7 171 4 US-09-621-976-7994 Sequence 7994, Ap  
115 13.8 65.7 293 4 US-09-313-294A-3503 Sequence 3503, Ap  
116 13.8 65.7 481 4 US-09-621-976-66 Sequence 66, Appli  
117 13.8 65.7 567 4 US-09-252-991A-12764 Sequence 12764, A  
118 13.8 65.7 693 4 US-09-252-991A-12990 Sequence 12990, A  
119 13.8 65.7 700 4 US-09-702-705-184 Sequence 184, App  
120 13.8 65.7 700 4 US-09-736-457-184 Sequence 184, App  
121 13.8 65.7 700 4 US-09-614-124B-184 Sequence 184, App  
122 13.8 65.7 700 4 US-09-671-325-184 Sequence 184, App  
123 13.8 65.7 700 4 US-09-589-184-184 Sequence 184, App  
124 13.8 65.7 735 4 US-09-252-991A-13191 Sequence 13191, A  
125 13.8 65.7 900 3 US-08-987-121A-1 Sequence 1, Appli  
126 13.8 65.7 900 3 US-08-987-121A-2 Sequence 2, Appli  
127 13.8 65.7 973 6 5212074-2 Patent No. 5212074  
128 13.8 65.7 1008 3 US-08-987-121A-5 Sequence 5, Appli  
129 13.8 65.7 1011 3 US-08-987-121A-3 Sequence 3, Appli  
130 13.8 65.7 1011 3 US-09-066-512-1 Sequence 1, Appli  
131 13.8 65.7 1095 4 US-09-328-352-533 Sequence 533, App  
132 13.8 65.7 1265 6 5198542-5 Patent No. 5198542  
133 13.8 65.7 1884 4 US-09-687-538B-5 Sequence 5, Appli  
134 13.8 65.7 1939 6 5198542-3 Patent No. 5198542  
135 13.8 65.7 2007 3 US-08-941-445A-8 Sequence 8, Appli  
136 13.8 65.7 2085 1 US-08-572-951-2 Sequence 2, Appli  
137 13.8 65.7 3468 4 US-09-221-017B-893 Sequence 893, App  
138 13.8 65.7 4065 6 5240838-14 Patent No. 5240838  
139 13.8 65.7 4379 1 US-08-592-214A-17 Sequence 17, Appli  
140 13.8 65.7 4379 3 US-09-149-976-17 Sequence 17, Appli  
141 13.8 65.7 17041 1 US-08-076-011-1 Sequence 1, Appli  
142 13.8 65.7 37030 4 US-08-311-731A-25 Sequence 25, Appli  
143 13.8 65.7 41171 4 US-08-311-731A-122 Sequence 122, App  
144 13.8 65.7 43795 3 US-08-742-185-101 Sequence 101, App  
145 13.8 65.7 193303 4 US-09-497-855A-37 Sequence 37, Appli  
146 13.8 65.7 193303 4 US-09-497-855A-44 Sequence 44, Appli  
147 13.6 64.8 88 4 US-09-465-355-29 Sequence 29, Appli  
148 13.6 64.8 189 4 US-09-543-681A-128 Sequence 128, App  
149 13.6 64.8 194 4 US-09-621-976-12549 Sequence 12549, A  
150 13.6 64.8 207 4 US-09-489-039A-3289 Sequence 3289, Ap  
151 13.6 64.8 261 4 US-09-389-681-411 Sequence 411, App  
152 13.6 64.8 261 4 US-09-620-403B-411 Sequence 411, App  
153 13.6 64.8 261 4 US-09-433-826B-411 Sequence 411, App  
154 13.6 64.8 261 4 US-09-604-287A-411 Sequence 411, App  
155 13.6 64.8 261 4 US-09-834-753-411 Sequence 411, App  
156 13.6 64.8 273 3 US-09-172-108-43 Sequence 43, Appli  
157 13.6 64.8 273 3 US-09-172-711-40 Sequence 40, Appli  
158 13.6 64.8 342 4 US-08-607-384A-4 Sequence 4, Appli  
159 13.6 64.8 348 4 US-08-651-155B-255 Sequence 255, App  
160 13.6 64.8 348 4 US-09-221-017B-764 Sequence 764, App  
161 13.6 64.8 348 4 US-09-194-036B-255 Sequence 255, App  
162 13.6 64.8 378 4 US-09-199-637A-316 Sequence 316, App  
163 13.6 64.8 393 4 US-09-313-294A-4256 Sequence 4256, Ap  
164 13.6 64.8 402 4 US-09-252-991A-7024 Sequence 7024, Ap  
165 13.6 64.8 513 4 US-09-621-976-613 Sequence 613, App  
166 13.6 64.8 598 3 US-08-998-416-1090 Sequence 1090, App  
167 13.6 64.8 686 3 US-08-998-416-1040 Sequence 1040, Ap  
168 13.6 64.8 720 1 US-08-044-621B-1 Sequence 1, Appli  
169 13.6 64.8 729 4 US-09-134-000C-2529 Sequence 2529, Ap  
170 13.6 64.8 741 4 US-09-489-039A-849 Sequence 849, App  
171 13.6 64.8 792 4 US-09-252-991A-7724 Sequence 7724, App  
172 13.6 64.8 819 4 US-09-199-637A-310 Sequence 310, App  
173 13.6 64.8 831 4 US-09-252-991A-7052 Sequence 7052, Ap

Sequence 1, Appli  
Patent No. 5206352  
Sequence 4, Appli  
Sequence 5, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 7994, Ap  
Sequence 3503, Ap  
Sequence 66, Appli  
Sequence 12764, A  
Sequence 12990, A  
Sequence 184, App  
Sequence 184, App  
Sequence 184, App  
Sequence 184, App  
Sequence 13191, A  
Sequence 1, Appli  
Sequence 2, Appli  
Patent No. 5212074  
Sequence 5, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 533, App  
Patent No. 5198542  
Sequence 5, Appli  
Patent No. 5198542  
Sequence 8, Appli  
Sequence 2, Appli  
Sequence 893, App  
Patent No. 5240838  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 1, Appli  
Sequence 25, Appli  
Sequence 122, App  
Sequence 101, App  
Sequence 37, Appli  
Sequence 44, Appli  
Sequence 29, Appli  
Sequence 128, App  
Sequence 12549, A  
Sequence 3289, Ap  
Sequence 411, App  
Sequence 411, App  
Sequence 411, App  
Sequence 411, App  
Sequence 43, Appli  
Sequence 40, Appli  
Sequence 4, Appli  
Sequence 255, App  
Sequence 764, App  
Sequence 255, App  
Sequence 316, App  
Sequence 4256, Ap  
Sequence 7024, Ap  
Sequence 613, App  
Sequence 1090, App  
Sequence 1040, Ap  
Sequence 1, Appli  
Sequence 2529, Ap  
Sequence 849, App  
Sequence 7724, App  
Sequence 310, App  
Sequence 7052, Ap

Sequence 55, Appli  
Sequence 40, Appli  
Sequence 314, App  
Sequence 4829, Ap  
Sequence 94, Appli  
Sequence 312, App  
Sequence 4861, Ap

174 13.6 64.8 990 4 US-09-205-258-55 Sequence 55, Appli  
175 13.6 64.8 1000 3 US-09-018-584A-40 Sequence 40, Appli  
176 13.6 64.8 1014 4 US-09-199-637A-314 Sequence 314, App  
177 13.6 64.8 1014 4 US-09-252-991A-4829 Sequence 4829, Ap  
178 13.6 64.8 1024 4 US-09-328-475C-94 Sequence 94, Appli  
179 13.6 64.8 1041 4 US-09-199-637A-312 Sequence 312, App  
180 13.6 64.8 1041 4 US-09-252-991A-4861 Sequence 4861, Ap

ALIGNMENTS

US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brum, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Eliens, Harma  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: PA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120.513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4233 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
US-09-120-513-1

Query Match 90.5%; Score 19; DB 3; Length 4233;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTGTGCTATGGCTCT 20  
Db 928 CCTGTGTGCTATGGCTCT 946

RESULT 2  
US-09-450-105-1  
; Sequence 1, Application US/09450105  
; Patent No. 6165166  
; GENERAL INFORMATION:



```
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Cheney
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1998-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

Query Match          90.5%; Score 19; DB 3; Length 4233;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCTGTTGGTCTATGCGTCT 20
Db 928 CCTGTTGGTCTATGCGTCT 946

RESULT 3
US-09-894-844-65
; Sequence 65, Application US/09894844
; Patent No. 6686166
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-65

Query Match          84.8%; Score 17.8; DB 4; Length 1860;
Best Local Similarity 90.5%; Pred. No. 6;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTGTTGGTCTATGCGTCTG 21
Db 1255 CCTGCTGCTCGATGCGTCTG 1275

RESULT 4
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
```

```
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          84.8%; Score 17.8; DB 3; Length 4403765;
Best Local Similarity 90.5%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTGTTGGTCTATGCGTCTG 21
Db 381864 CCTGCTGCTCGATGCGTCTG 381884

RESULT 5
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match          84.8%; Score 17.8; DB 3; Length 4411529;
Best Local Similarity 90.5%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTGTTGGTCTATGCGTCTG 21
Db 381808 CCTGCTGCTCGATGCGTCTG 381828

RESULT 6
US-08-276-967-1
; Sequence 1, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
```

```

; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: US/418\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7785 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-276-967-1

```

```

Query Match          77.1%; Score 16.2; DB 2; Length 7785;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 CCCGTGCTATGCGTCTG 21
Db 530 CCCCTATGCTATGCGTCTG 550

```

```

RESULT 7
US-09-801-876B-3
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3

```

```

Query Match          76.2%; Score 16; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 6 TTGGTCTATGCGTCTG 21
Db 141042 TTGGTCTATGCGTCTG 141057

```

```

RESULT 8
US-10-254-869-3
; Sequence 3, Application US/10254869

```

```

; Patent No. 6653117
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160DIV
; CURRENT APPLICATION NUMBER: US/10/254,869
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

```

```

Query Match          76.2%; Score 16; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 6 TTGGTCTATGCGTCTG 21
Db 141042 TTGGTCTATGCGTCTG 141057

```

```

RESULT 9
US-09-741-150-3/c
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

```

```

Query Match          75.2%; Score 15.8; DB 4; Length 112132;
Best Local Similarity 89.5%; Pred. No. 97;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 3 CTGTTGGTCTATGCGTCTG 21
Db 72921 CTATGGTCTATGCTCTG 72903

```

```

RESULT 10
US-10-160-187-3/c
; Sequence 3, Application US/10160187
; Patent No. 6620607
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968DIV

```

; TYPE: DNA

1.15  
1.3  
1.5  
1.7  
1.9  
2.1  
2.3  
2.5  
2.7  
2.9  
3.1  
3.3  
3.5  
3.7  
3.9  
4.1  
4.3  
4.5  
4.7  
4.9  
5.1  
5.3  
5.5  
5.7  
5.9  
6.1  
6.3  
6.5  
6.7  
6.9  
7.1  
7.3  
7.5  
7.7  
7.9  
8.1  
8.3  
8.5  
8.7  
8.9  
9.1  
9.3  
9.5  
9.7  
9.9  
10.1  
10.3  
10.5  
10.7  
10.9  
11.1  
11.3  
11.5  
11.7  
11.9  
12.1  
12.3  
12.5  
12.7  
12.9  
13.1  
13.3  
13.5  
13.7  
13.9  
14.1  
14.3  
14.5  
14.7  
14.9  
15.1  
15.3  
15.5  
15.7  
15.9  
16.1  
16.3  
16.5  
16.7  
16.9  
17.1  
17.3  
17.5  
17.7  
17.9  
18.1  
18.3  
18.5  
18.7  
18.9  
19.1  
19.3  
19.5  
19.7  
19.9  
20.1  
20.3  
20.5  
20.7  
20.9  
21.1  
21.3  
21.5  
21.7  
21.9  
22.1  
22.3  
22.5  
22.7  
22.9  
23.1  
23.3  
23.5  
23.7  
23.9  
24.1  
24.3  
24.5  
24.7  
24.9  
25.1  
25.3  
25.5  
25.7  
25.9  
26.1  
26.3  
26.5  
26.7  
26.9  
27.1  
27.3  
27.5  
27.7  
27.9  
28.1  
28.3  
28.5  
28.7  
28.9  
29.1  
29.3  
29.5  
29.7  
29.9  
30.1  
30.3  
30.5  
30.7  
30.9  
31.1  
31.3  
31.5  
31.7  
31.9  
32.1  
32.3  
32.5  
32.7  
32.9  
33.1  
33.3  
33.5  
33.7  
33.9  
34.1  
34.3  
34.5  
34.7  
34.9  
35.1  
35.3  
35.5  
35.7  
35.9  
36.1  
36.3  
36.5  
36.7  
36.9  
37.1  
37.3  
37.5  
37.7  
37.9  
38.1  
38.3  
38.5  
38.7  
38.9  
39.1  
39.3  
39.5  
39.7  
39.9  
40.1  
40.3  
40.5  
40.7  
40.9  
41.1  
41.3  
41.5  
41.7  
41.9  
42.1  
42.3  
42.5  
42.7  
42.9  
43.1  
43.3  
43.5  
43.7  
43.9  
44.1  
44.3  
44.5  
44.7  
44.9  
45.1  
45.3  
45.5  
45.7  
45.9  
46.1  
46.3  
46.5  
46.7  
46.9  
47.1  
47.3  
47.5  
47.7  
47.9  
48.1  
48.3  
48.5  
48.7  
48.9  
49.1  
49.3  
49.5  
49.7  
49.9  
50.1  
50.3  
50.5  
50.7  
50.9  
51.1  
51.3  
51.5  
51.7  
51.9  
52.1  
52.3  
52.5  
52.7  
52.9  
53.1  
53.3  
53.5  
53.7  
53.9  
54.1  
54.3  
54.5  
54.7  
54.9  
55.1  
55.3  
55.5  
55.7  
55.9  
56.1  
56.3  
56.5  
56.7  
56.9  
57.1  
57.3  
57.5  
57.7  
57.9  
58.1  
58.3  
58.5  
58.7  
58.9  
59.1  
59.3  
59.5  
59.7  
59.9  
60.1  
60.3  
60.5  
60.7  
60.9  
61.1  
61.3  
61.5  
61.7  
61.9  
62.1  
62.3  
62.5  
62.7  
62.9  
63.1  
63.3  
63.5  
63.7  
63.9  
64.1  
64.3  
64.5  
64.7  
64.9  
65.1  
65.3  
65.5  
65.7  
65.9  
66.1  
66.3  
66.5  
66.7  
66.9  
67.1  
67.3  
67.5  
67.7  
67.9  
68.1  
68.3  
68.5  
68.7  
68.9  
69.1  
69.3  
69.5  
69.7  
69.9  
70.1  
70.3  
70.5  
70.7  
70.9  
71.1  
71.3  
71.5  
71.7  
71.9  
72.1  
72.3  
72.5  
72.7  
72.9  
73.1  
73.3  
73.5  
73.7  
73.9  
74.1  
74.3  
74.5  
74.7  
74.9  
75.1  
75.3  
75.5  
75.7  
75.9  
76.1  
76.3  
76.5  
76.7  
76.9  
77.1  
77.3  
77.5  
77.7  
77.9  
78.1  
78.3  
78.5  
78.7  
78.9  
79.1  
79.3  
79.5  
79.7  
79.9  
80.1  
80.3  
80.5  
80.7  
80.9  
81.1  
81.3  
81.5  
81.7  
81.9  
82.1  
82.3  
82.5  
82.7  
82.9  
83.1  
83.3  
83.5  
83.7  
83.9  
84.1  
84.3  
84.5  
84.7  
84.9  
85.1  
85.3  
85.5  
85.7  
85.9  
86.1  
86.3  
86.5  
86.7  
86.9  
87.1  
87.3  
87.5  
87.7  
87.9  
88.1  
88.3  
88.5  
88.7  
88.9  
89.1  
89.3  
89.5  
89.7  
89.9  
90.1  
90.3  
90.5  
90.7  
90.9  
91.1  
91.3  
91.5  
91.7  
91.9  
92.1  
92.3  
92.5  
92.7  
92.9  
93.1  
93.3  
93.5  
93.7  
93.9  
94.1  
94.3  
94.5  
94.7  
94.9  
95.1  
95.3  
95.5  
95.7  
95.9  
96.1  
96.3  
96.5  
96.7  
96.9  
97.1  
97.3  
97.5  
97.7  
97.9  
98.1  
98.3  
98.5  
98.7  
98.9  
99.1  
99.3  
99.5  
99.7  
99.9  
100.1  
100.3  
100.5  
100.7  
100.9  
101.1  
101.3  
101.5  
101.7  
101.9  
102.1  
102.3  
102.5  
102.7  
102.9  
103.1  
103.3  
103.5  
103.7  
103.9  
104.1  
104.3  
104.5  
104.7  
104.9  
105.1  
105.3  
105.5  
105.7  
105.9  
106.1  
106.3  
106.5  
106.7  
106.9  
107.1  
107.3  
107.5  
107.7  
107.9  
108.1  
108.3  
108.5  
108.7  
108.9  
109.1  
109.3  
109.5  
109.7  
109.9  
110.1  
110.3  
110.5  
110.7  
110.9  
111.1  
111.3  
111.5  
111.7  
111.9  
112.1  
112.3  
112.5  
112.7  
112.9  
113.1  
113.3  
113.5  
113.7  
113.9  
114.1  
114.3  
114.5  
114.7  
114.9  
115.1  
115.3  
115.5  
115.7  
115.9  
116.1  
116.3  
116.5  
116.7  
116.9  
117.1  
117.3  
117.5  
117.7  
117.9  
118.1  
118.3  
118.5  
118.7  
118.9  
119.1  
119.3  
119.5  
119.7  
119.9  
120.1  
120.3  
120.5  
120.7  
120.9  
121.1  
121.3  
121.5  
121.7  
121.9  
122.1  
122.3  
122.5  
122.7  
122.9  
123.1  
123.3  
123.5  
123.7  
123.9  
124.1  
124.3  
124.5  
124.7  
124.9  
125.1  
125.3  
125.5  
125.7  
125.9  
126.1  
126.3  
126.5  
126.7  
126.9  
127.1  
127.3  
127.5  
127.7  
127.9  
128.1  
128.3  
128.5  
128.7  
128.9  
129.1  
129.3  
129.5  
129.7  
129.9  
130.1  
130.3  
130.5  
130.7  
130.9  
131.1  
131.3  
131.5  
131.7  
131.9  
132.1  
132.3  
132.5  
132.7  
132.9  
133.1  
133.3

4 TGTGGTCTATGCGTCTG 21

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-418-444A-1

Query Match          70.5%; Score 14.8; DB 1; Length 3088;
Best Local Similarity 88.9%; Pred.No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGTGTGCTATGCGTCTG 21
    ||||| ||||| |||||
Db 2961 TGTGTGCTATGCGTCTG 2978

RESULT 17
US-08-961-527-54
; Sequence 54, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-54

Query Match          70.5%; Score 14.8; DB 4; Length 20986;
Best Local Similarity 88.9%; Pred.No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTGTGTGCTATGCGTCTG 19
    ||||| ||||| |||||
Db 4404 CCTGTGTGCTATGCGTCTG 4421

RESULT 18
US-08-311-731A-136/c
; Sequence 136, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-418-444A-1

Query Match          70.5%; Score 14.8; DB 1; Length 3088;
Best Local Similarity 88.9%; Pred.No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGTGTGCTATGCGTCTG 21
    ||||| ||||| |||||
Db 2961 TGTGTGCTATGCGTCTG 2978

RESULT 17
US-08-961-527-54
; Sequence 54, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-54

Query Match          70.5%; Score 14.8; DB 4; Length 20986;
Best Local Similarity 88.9%; Pred.No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTGTGTGCTATGCGTCTG 19
    ||||| ||||| |||||
Db 4404 CCTGTGTGCTATGCGTCTG 4421

RESULT 18
US-08-311-731A-136/c
; Sequence 136, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
```

APPLICANT: SMITH, DOUGLAS  
APPLICANT: MAO, JEN-I  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 411  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,731A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: C0044/7125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
INFORMATION FOR SEQ ID NO: 136:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36138 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mycobacterium leprae  
US-08-311-731A-136  
Query Match 70.5%; Score 14.8; DB 4; Length 36138;  
Best Local Similarity 88.9%; Pred. No. 2.7e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CCCTGTTGGTCTATGCGT 18  
Db 21153 CCCTGTTGGTCTATGCGT 21136  
RESULT 19  
US-09-313-294A-33/c  
Sequence 33, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 33  
LENGTH: 271  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: lncyte ID No. 6476212 700548369H1  
US-09-313-294A-33

Query Match 69.5%; Score 14.6; DB 4; Length 271;  
Best Local Similarity 81.0%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
Db 101 CACTGTTGGTCACTGCGTCTG 81  
RESULT 20  
US-09-280-116-164/c  
Sequence 164, Application US/09280116A  
Patent No. 6331427  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
FILE REFERENCE: 5800-24, 035800/176965  
CURRENT APPLICATION NUMBER: US/09/280,116A  
CURRENT FILING DATE: 1999-03-26  
NUMBER OF SEQ ID NOS: 268  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 164  
LENGTH: 442  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: cathepsin  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(442)  
OTHER INFORMATION: n = a, t, c or g  
US-09-280-116-164

Query Match 69.5%; Score 14.6; DB 4; Length 442;  
Best Local Similarity 81.0%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
Db 81 CCCTGTTGGTCTATGCGTCTG 61

RESULT 21  
US-09-621-976-17799/c  
Sequence 17799, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 17799  
LENGTH: 525  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-17799

Query Match 69.5%; Score 14.6; DB 4; Length 525;  
Best Local Similarity 81.0%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
Db 45 CCACATTGGTCCATGCGTCTG 25

RESULT 22

```
US-09-252-991A-10362
; Sequence 10362, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10362
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10362

Query Match          69.5%; Score 14.6; DB 4; Length 576;
Best Local Similarity 81.0%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 409 CCCTGCGGTCGATCGGTTG 429

RESULT 23
US-08-466-033-116
; Sequence 116, Application US/08466033
; Patent No. 5766840
; GENERAL INFORMATION:
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Wages, John
; APPLICANT: Young, LaVonne M.
; APPLICANT: Fry, Kirk E.
; APPLICANT: Linnen, Jeffrey M.
; TITLE OF INVENTION: Hepatitis G Virus and Molecular
; TITLE OF INVENTION: Cloning Thereof
; NUMBER OF SEQUENCES: 277
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,033
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,886
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,509
; FILING DATE: 16-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,729
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,271
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,558
```

```
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA: US 08/285,543
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,985
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4603-0201.36/G100P11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Consensus Sequence 11E3
; Patent No. 5766840
US-08-466-033-116

Query Match          69.5%; Score 14.6; DB 1; Length 590;
Best Local Similarity 81.0%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
Db 69 CCAGTATGCTATGGGTCG 89

RESULT 24
US-08-444-733-116
; Sequence 116, Application US/08444733
; Patent No. 5824507
; GENERAL INFORMATION:
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Wages, John
; APPLICANT: Young, LaVonne M.
; APPLICANT: Fry, Kirk E.
; APPLICANT: Linnen, Jeffrey M.
; TITLE OF INVENTION: Hepatitis G Virus and Molecular
; TITLE OF INVENTION: Cloning Thereof
; NUMBER OF SEQUENCES: 277
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,733
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,886
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,509
; FILING DATE: 16-DEC-1994
; PRIOR APPLICATION DATA:
```

GenCore version 5.1.6  
Copyright (c) 1993 -2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:48:43 ; Search time 128.841 Seconds  
(without alignments) 729.318 Million cell updates

Title: US-09-877-819E-40

Perfect score: 21  
Sequence: 1 ccctattggtctatgcgtctg 21

scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters:

Minimum DB sec length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Listing first 180 summaries

Database : Published\_Applications\_NA:\*

```

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US05_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10C_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	10	US-09-877-8193-40
2	21	100.0	254	10	US-09-877-8193-55
3	21	100.0	267	9	US-09-796-692-143
4	21	100.0	267	9	US-09-796-692-328
5	21	100.0	267	9	US-09-796-692-328
6	21	100.0	267	9	US-09-796-692-7075
7	21	100.0	267	15	US-10-040-862-143
8	21	100.0	267	15	US-10-040-862-328
9	21	100.0	267	15	US-10-040-862-7075
10	21	100.0	267	16	US-10-057-4755-143
11	21	100.0	267	16	US-10-057-4755-328
12	21	100.0	267	16	US-10-057-4755-7075
13	21	100.0	267	16	US-10-154-8848-143
14	21	100.0	267	16	US-10-154-8848-328
15	21	100.0	267	16	US-10-154-8848-7075
16	21	100.0	267	16	US-10-154-8848-7075
17	21	100.0	267	16	US-10-154-8848-7075
18	21	100.0	267	16	US-10-154-8848-7075
19	21	100.0	267	16	US-10-154-8848-7075
20	21	100.0	267	16	US-10-154-8848-7075
21	21	100.0	267	16	US-10-154-8848-7075
22	21	100.0	267	16	US-10-154-8848-7075
23	21	100.0	267	16	US-10-154-8848-7075
24	21	100.0	267	16	US-10-154-8848-7075
25	21	100.0	267	16	US-10-154-8848-7075
26	21	100.0	267	16	US-10-154-8848-7075
27	21	100.0	267	16	US-10-154-8848-7075
28	21	100.0	267	16	US-10-154-8848-7075
29	21	100.0	267	16	US-10-154-8848-7075
30	21	100.0	267	16	US-10-154-8848-7075
31	21	100.0	267	16	US-10-154-8848-7075
32	21	100.0	267	16	US-10-154-8848-7075
33	21	100.0	267	16	US-10-154-8848-7075
34	21	100.0	267	16	US-10-154-8848-7075
35	21	100.0	267	16	US-10-154-8848-7075
36	21	100.0	267	16	US-10-154-8848-7075
37	21	100.0	267	16	US-10-154-8848-7075
38	21	100.0	267	16	US-10-154-8848-7075
39	21	100.0	267	16	US-10-154-8848-7075
40	21	100.0	267	16	US-10-154-8848-7075
41	21	100.0	267	16	US-10-154-8848-7075
42	21	100.0	267	16	US-10-154-8848-7075
43	21	100.0	267	16	US-10-154-8848-7075
44	21	100.0	267	16	US-10-154-8848-7075
45	21	100.0	267	16	US-10-154-8848-7075
46	21	100.0	267	16	US-10-154-8848-7075
47	21	100.0	267	16	US-10-154-8848-7075
48	21	100.0	267	16	US-10-154-8848-7075
49	21	100.0	267	16	US-10-154-8848-7075
50	21	100.0	267	16	US-10-154-8848-7075
51	21	100.0	267	16	US-10-154-8848-7075
52	21	100.0	267	16	US-10-154-8848-7075
53	21	100.0	267	16	US-10-154-8848-7075
54	21	100.0	267	16	US-10-154-8848-7075
55	21	100.0	267	16	US-10-154-8848-7075
56	21	100.0	267	16	US-10-154-8848-7075
57	21	100.0	267	16	US-10-154-8848-7075
58	21	100.0	267	16	US-10-154-8848-7075
59	21	100.0	267	16	US-10-154-8848-7075
60	21	100.0	267	16	US-10-154-8848-7075
61	21	100.0	267	16	US-10-154-8848-7075
62	21	100.0	267	16	US-10-154-8848-707

```

c 88 15.8 75.2 550 13 US-10-027-632-322684 Sequence 322684, 161 15.2 72.4 643 13 US-10-027-632-179884 Sequence 179884,
c 89 15.8 75.2 550 13 US-10-027-632-322685 Sequence 322685, 162 15.2 72.4 643 16 US-10-027-632-179884 Sequence 179884,
c 90 15.8 75.2 550 16 US-10-027-632-302207 Sequence 302207, 163 15.2 72.4 860 13 US-10-027-632-139212 Sequence 139212,
c 91 15.8 75.2 550 16 US-10-027-632-302208 Sequence 302208, 164 15.2 72.4 860 16 US-10-027-632-139212 Sequence 139212,
c 92 15.8 75.2 550 16 US-10-027-632-322684 Sequence 322684, c 165 15.2 72.4 1043 13 US-10-027-632-257024 Sequence 257024,
c 93 15.8 75.2 550 16 US-10-027-632-322685 Sequence 322685, c 166 15.2 72.4 1043 16 US-10-027-632-257024 Sequence 257024,
c 94 15.8 75.2 562 13 US-10-027-632-44510 Sequence 44510, c 167 15.2 72.4 1178 9 US-09-924-256A-93 Sequence 93, Appl
c 95 15.8 75.2 562 13 US-10-027-632-301126 Sequence 301126, 168 15.2 72.4 1181 13 US-10-027-632-313843 Sequence 313843,
c 96 15.8 75.2 562 16 US-10-027-632-44510 Sequence 44510, 169 15.2 72.4 1181 16 US-10-027-632-313843 Sequence 313843,
c 97 15.8 75.2 562 16 US-10-027-632-301126 Sequence 301126, 170 15.2 72.4 1271 13 US-10-027-632-211382 Sequence 211382,
c 98 15.8 75.2 564 13 US-10-027-632-35201 Sequence 35201, A 171 15.2 72.4 1271 16 US-10-027-632-211382 Sequence 211382,
c 99 15.8 75.2 564 13 US-10-027-632-35202 Sequence 35202, A 172 15.2 72.4 1554 9 US-09-738-626-2728 Sequence 2728, Ap
c 100 15.8 75.2 564 16 US-10-027-632-35201 Sequence 35201, A 173 15.2 72.4 1575 13 US-10-425-114-31872 Sequence 31872, A
c 101 15.8 75.2 564 16 US-10-027-632-35202 Sequence 35202, A 174 15.2 72.4 2338 15 US-10-067-449-12 Sequence 12, Appl
c 102 15.8 75.2 582 13 US-10-027-632-238862 Sequence 238862, 175 15.2 72.4 2338 15 US-10-067-449-12 Sequence 12, Appl
c 103 15.8 75.2 582 13 US-10-027-632-238863 Sequence 238863, 176 15.2 72.4 2338 15 US-10-067-449-12 Sequence 12, Appl
c 104 15.8 75.2 582 16 US-10-027-632-238864 Sequence 238864, c 177 15.2 72.4 2859 13 US-10-027-632-112342 Sequence 112342,
c 105 15.8 75.2 582 16 US-10-027-632-238862 Sequence 238862, c 178 15.2 72.4 2859 16 US-10-027-632-112342 Sequence 112342,
c 106 15.8 75.2 582 16 US-10-027-632-238863 Sequence 238863, c 179 15.2 72.4 3175 16 US-10-074-024-812 Sequence 812, App
c 107 15.8 75.2 582 16 US-10-027-632-238864 Sequence 238864, 180 15.2 72.4 3175 16 US-10-242-355-934 Sequence 934, App
c 108 15.8 75.2 585 13 US-10-027-632-79299 Sequence 79299, A
c 109 15.8 75.2 585 16 US-10-027-632-79299 Sequence 79299, A
c 110 15.8 75.2 611 13 US-10-027-632-220215 Sequence 220215,
c 111 15.8 75.2 611 13 US-10-027-632-220216 Sequence 220216,
c 112 15.8 75.2 611 13 US-10-027-632-220217 Sequence 220217,
c 113 15.8 75.2 611 16 US-10-027-632-220215 Sequence 220215,
c 114 15.8 75.2 611 16 US-10-027-632-220216 Sequence 220216,
c 115 15.8 75.2 611 16 US-10-027-632-220217 Sequence 220217,
c 116 15.8 75.2 615 13 US-10-027-632-214892 Sequence 214892,
c 117 15.8 75.2 615 16 US-10-027-632-214892 Sequence 214892,
c 118 15.8 75.2 1086 16 US-10-369-493-26312 Sequence 26312, A
c 119 15.8 75.2 1329 13 US-10-424-598-6159 Sequence 6159, Ap
c 120 15.8 75.2 108162 13 US-10-085-783A-57491 Sequence 57491, A
c 121 15.8 75.2 194049 13 US-10-087-132-1495 Sequence 1495, Ap
c 122 15.8 75.2 302603 13 US-10-271-416-8 Sequence 8, Appl
c 123 15.8 75.2 322101 13 US-10-354-247-1 Sequence 1, Appl
c 124 15.8 75.2 322101 15 US-10-060-902-1 Sequence 1, Appl
c 125 15.4 73.3 465 13 US-10-085-783A-57491 Sequence 57491, A
c 126 15.4 73.3 465 16 US-10-424-598-6159 Sequence 6159, A
c 127 15.4 73.3 1257 13 US-10-027-632-124090 Sequence 124090,
c 128 15.4 73.3 1257 16 US-10-027-632-124090 Sequence 124090,
c 129 15.4 73.3 1322 13 US-10-425-114-36038 Sequence 36038, A
c 130 15.4 73.3 1353 13 US-10-282-132A-11817 Sequence 11817, A
c 131 15.4 73.3 172637 9 US-09-805-458A-3 Sequence 3, Appl
c 132 15.2 72.4 128 13 US-10-424-598-92120 Sequence 92120, A
c 133 15.2 72.4 140 9 US-09-864-761-26779 Sequence 26779, A
c 134 15.2 72.4 140 15 US-10-029-386-18342 Sequence 18342, A
c 135 15.2 72.4 302 13 US-10-424-598-59675 Sequence 59675, A
c 136 15.2 72.4 332 13 US-10-085-783A-37520 Sequence 37520, A
c 137 15.2 72.4 332 16 US-10-242-535A-37520 Sequence 37520, A
c 138 15.2 72.4 395 9 US-09-954-456-1864 Sequence 1864, Ap
c 139 15.2 72.4 401 13 US-10-027-632-35368 Sequence 35368, A
c 140 15.2 72.4 401 13 US-10-027-632-62406 Sequence 62406, A
c 141 15.2 72.4 401 13 US-10-027-632-63010 Sequence 63010, A
c 142 15.2 72.4 401 16 US-10-027-632-35368 Sequence 35368, A
c 143 15.2 72.4 401 16 US-10-027-632-62406 Sequence 62406, A
c 144 15.2 72.4 401 16 US-10-027-632-63010 Sequence 63010, A
c 145 15.2 72.4 439 9 US-09-880-107-1855 Sequence 1855, Ap
c 146 15.2 72.4 446 9 US-09-867-701-1919 Sequence 1919, Ap
c 147 15.2 72.4 477 9 US-09-864-761-10138 Sequence 10138, A
c 148 15.2 72.4 481 13 US-10-027-632-306533 Sequence 306533,
c 149 15.2 72.4 485 16 US-10-027-632-306533 Sequence 306533,
c 150 15.2 72.4 485 15 US-10-198-845-54 Sequence 54, Appl
c 151 15.2 72.4 556 15 US-10-029-386-4586 Sequence 4586, Ap
c 152 15.2 72.4 590 13 US-10-027-632-76802 Sequence 76802, A
c 153 15.2 72.4 590 13 US-10-027-632-76803 Sequence 76803, A
c 154 15.2 72.4 590 16 US-10-027-632-76802 Sequence 76802, A
c 155 15.2 72.4 590 16 US-10-027-632-76803 Sequence 76803, A
c 156 15.2 72.4 629 13 US-10-027-632-33032 Sequence 33032, A
c 157 15.2 72.4 629 16 US-10-027-632-33032 Sequence 33032, A
c 158 15.2 72.4 630 16 US-10-264-048-278 Sequence 278, App
c 159 15.2 72.4 635 13 US-10-027-632-211383 Sequence 211383,
c 160 15.2 72.4 635 16 US-10-027-632-211383 Sequence 211383,

```

## ALIGNMENTS

## RESULT 1

US-09-877-819B-40

; Sequence 40, Application US/09877819B

; Publication No. US20030190609A1

; GENERAL INFORMATION:

; APPLICANT: Torney, David

; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing

; FILE REFERENCE: S-94,864

; CURRENT APPLICATION NUMBER: US/09/877,819B

; CURRENT FILING DATE: 2001-06-07

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 40

; LENGTH: 21

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Address sequence

US-09-877-819B-40

Query Match 100.0%; Score 21; DB 10; Length 21;

Best Local Similarity 100.0%; Pred.No. 0.55;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTTGGTCTATGCGTCTG 21

Db 1 CCTCTTGGTCTATGCGTCTG 21

## RESULT 2

US-09-877-819B-55/c

; Sequence 55, Application US/09877819B

; Publication No. US20030190609A1

; GENERAL INFORMATION:

; APPLICANT: Torney, David

; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing

; FILE REFERENCE: S-94,864

; CURRENT APPLICATION NUMBER: US/09/877,819B

; CURRENT FILING DATE: 2001-06-07

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 55

; LENGTH: 254

; TYPE: DNA

; ORGANISM: Human HLA

US-09-877-819B-55



```
Query Match      100.0%; Score 21; DB 10; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21
Db 60 CCCTGTTGGTCTATGCGCTG 40

RESULT 3
US-09-796-692-143/c
; Sequence 143, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-143

Query Match      100.0%; Score 21; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21
Db 181 CCCTGTTGGTCTATGCGCTG 161

RESULT 4
US-09-796-692-328/c
; Sequence 328, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
```

```
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-328

Query Match      100.0%; Score 21; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21
Db 181 CCCTGTTGGTCTATGCGCTG 161

RESULT 5
US-09-796-692-7075/c
; Sequence 7075, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
```



```

; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 69/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7075

Query Match 100.0%; Score 21; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGTTGGTCTATGGGCTG 21
Db 181 CCCTGTTGGTCTATGGGCTG 161

RESULT 9
US-10-057-475B-143/c
; Sequence 143, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22

```

; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 328  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-057-475B-328

Query Match 100.0%; Score 21; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.66; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

Qy 1 CCCTGTTGGTCTATCGCTG 21  
|||  
Db 181 CCCTGTTGGTCTATCGCTG 161

RESULT 11  
US-10-057-475B-7075/c  
; Sequence 7075, Application US/10057475B  
; Publication No. US20040002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Clapper, Jonathan David  
; APPLICANT: Wang, Aijun  
; APPLICANT: Ordonez, Nadia  
; APPLICANT: Carter, Lauren  
; APPLICANT: McNeill, Patricia Dianne  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-014402US  
; CURRENT APPLICATION NUMBER: US/10/057,475B  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7075  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-057-475B-7075

Query Match 100.0%; Score 21; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.66; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

Qy 1 CCCTGTTGGTCTATCGCTG 21  
|||  
Db 181 CCCTGTTGGTCTATCGCTG 161

## RESULT 12

US-10-154-884B-143/c  
; Sequence 143, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 143  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-143

Query Match 100.0%; Score 21; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.66; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

Qy 1 CCCTGTTGGTCTATCGCTG 21  
|||  
Db 181 CCCTGTTGGTCTATCGCTG 161

## RESULT 13

US-10-154-884B-328/c  
; Sequence 328, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B

; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 328  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-328

Query Match 100.0%; Score 21; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||||  
Db 181 CCCTGTTGGTCTATGCGCTG 161

RESULT 14  
US-10-154-884B-7075/c  
; Sequence 7075, Application US/101548843  
; Publication No. US20040003561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903

; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 7075  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-7075

Query Match 100.0%; Score 21; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||||  
Db 181 CCCTGTTGGTCTATGCGCTG 161

RESULT 15  
US-09-796-692-42/c  
; Sequence 42, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; NUMBER OF SEQ ID NOS: 9537  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 42  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-42

Query Match 100.0%; Score 21; DB 9; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||||  
Db 196 CCCTGTTGGTCTATGCGCTG 166

```
RESULT 16
US-09-796-692-3554/c
; Sequence 3554, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Mannion, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077-001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-3554

Query Match 100.0%; Score 21; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGCTCTG 21
Db 186 CCCTGTTGGTCTATCGCTCTG 166

RESULT 17
US-10-040-862-42/c
; Sequence 42, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Mannion, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-3554

Query Match 100.0%; Score 21; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGCTCTG 21
Db 186 CCCTGTTGGTCTATCGCTCTG 166

RESULT 18
US-10-040-862-3554/c
; Sequence 3554, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Mannion, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
```

;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3554  
;; LENGTH: 272  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-3554

Query Match 100.0%; Score 21; DB 15; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 186 CCTGTGGTCTATGCGTCTG 166

## RESULT 19

US-10-057-475B-42/c  
;; Sequence 42, Application US/10057475B  
;; Publication No. US20040002068A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Clapper, Jonathan David  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Ordonez, Nadia  
;; APPLICANT: Carter, Lauren  
;; APPLICANT: McNeill, Patricia Dianne  
;; APPLICANT: Corixa Corporation  
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
;; FILE REFERENCE: 014058-014402US  
;; CURRENT APPLICATION NUMBER: US/10/057,475B  
;; CURRENT FILING DATE: 2002-01-22  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/219,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 10979  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 42  
;; LENGTH: 272  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-057-475B-42

Query Match 100.0%; Score 21; DB 16; Length 272;

Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 186 CCTGTGGTCTATGCGTCTG 166

## RESULT 20

US-10-057-475B-3554/c  
;; Sequence 3554, Application US/10057475B  
;; Publication No. US20040002068A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Clapper, Jonathan David  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Ordonez, Nadia  
;; APPLICANT: Carter, Lauren  
;; APPLICANT: McNeill, Patricia Dianne  
;; APPLICANT: Corixa Corporation  
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
;; FILE REFERENCE: 014058-014402US  
;; CURRENT APPLICATION NUMBER: US/10/057,475B  
;; CURRENT FILING DATE: 2002-01-22  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/219,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 10979  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3554  
;; LENGTH: 272  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-057-475B-3554

Query Match 100.0%; Score 21; DB 16; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTGGTCTATGCGTCTG 21  
|||||  
Db 186 CCTGTGGTCTATGCGTCTG 166

## RESULT 21

US-10-154-884B-42/c  
;; Sequence 42, Application US/10154884B  
;; Publication No. US20040005561A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane

```
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-42
```

```
Query Match 100.0%; Score 21; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 186 CCCTGTTGGTCTATGCGTCTG 166

RESULT 22
US-10-154-884B-3554/C
; Sequence 3554, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
```

```
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-3554
```

```
Query Match 100.0%; Score 21; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CCCTGTTGGTCTATGCGTCTG 21
|||||
Db 186 CCCTGTTGGTCTATGCGTCTG 166
```

```
RESULT 23
US-09-796-692-280
; Sequence 280, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-280
```

```
Query Match 100.0%; Score 21; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



OY 1 CCCTGTTGGTCTATCGCTG 21  
Db 87 CCCTGTTGGTCTATCGCTG 107

## RESULT 24

US-09-796-692-530  
; Sequence 530, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 530  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-530

Query Match 100.0%; Score 21; DB 9; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCCTGTTGGTCTATCGCTG 21  
Db 87 CCCTGTTGGTCTATCGCTG 107

## RESULT 25

US-09-796-692-3270  
; Sequence 3270, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09796,692  
; CURRENT FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 3270  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-3270

Query Match 100.0%; Score 21; DB 9; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCCTGTTGGTCTATCGCTG 21  
Db 87 CCCTGTTGGTCTATCGCTG 107

## RESULT 26

US-10-040-862-280  
; Sequence 280, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950





;  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3270  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-057-475B-3270

Query Match 100.0%; Score 21; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
|||||

Db 87 CCCTGTTGGTCTATGCGTCTG 107

## RESULT 32

US-10-154-884B-280  
; Sequence 280, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23

; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 280  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-154-884B-280

Query Match 100.0%; Score 21; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
|||||

Db 87 CCCTGTTGGTCTATGCGTCTG 107

## RESULT 33

US-10-154-884B-530  
; Sequence 530, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23

; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 530  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-154-884B-530

Query Match 100.0%; Score 21; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
|||||

Db 87 CCCTGTTGGTCTATGCGTCTG 107

## RESULT 34

US-10-154-884B-3270  
; Sequence 3270, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23

; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17

;/ PRIOR APPLICATION NUMBER: US 60/200,545  
;/ PRIOR FILING DATE: 2000-04-27  
;/ PRIOR APPLICATION NUMBER: US 60/200,303  
;/ PRIOR FILING DATE: 2000-04-28  
;/ PRIOR APPLICATION NUMBER: US 60/200,779  
;/ PRIOR FILING DATE: 2000-04-28  
;/ PRIOR APPLICATION NUMBER: US 60/200,999  
;/ PRIOR FILING DATE: 2000-05-01  
;/ PRIOR APPLICATION NUMBER: US 60/202,054  
;/ PRIOR FILING DATE: 2000-05-04  
;/ PRIOR APPLICATION NUMBER: US 60/206,201  
;/ PRIOR FILING DATE: 2000-05-22  
;/ PRIOR APPLICATION NUMBER: US 60/218,950  
;/ PRIOR FILING DATE: 2000-07-14  
;/ PRIOR APPLICATION NUMBER: US 60/222,903  
;/ PRIOR FILING DATE: 2000-08-03  
;/ Remaining Prior Application data removed - See File Wrapper or PALM.  
;/ NUMBER OF SEQ ID NOS: 11290  
;/ SOFTWARE: FastSeq for Windows Version 3.0  
;/ SEQ ID NO 3270  
;/ LENGTH: 294  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
US-10-154-884B-3270

Query Match 100.0%; Score 21; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||  
Db 87 CCCTGTTGGTCTATGCGCTG 107

RESULT 35  
US-09-918-995-36460/c  
;/ Sequence 36460, Application US/09918995  
;/ Publication No. US20030073623A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Hyseq, Inc.  
;/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
;/ FILE REFERENCE: 20411-756  
;/ CURRENT APPLICATION NUMBER: US/09/918,995  
;/ CURRENT FILING DATE: 2001-07-30  
;/ PRIOR APPLICATION NUMBER: US/09/235,076  
;/ PRIOR FILING DATE: 1999-01-20  
;/ NUMBER OF SEQ ID NOS: 38054  
;/ SOFTWARE: FastSeq for Windows Version 3.0  
;/ SEQ ID NO 36460  
;/ LENGTH: 410  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
US-09-918-995-36460

Query Match 100.0%; Score 21; DB 10; Length 410;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||  
Db 204 CCCTGTTGGTCTATGCGCTG 184

RESULT 36  
US-09-918-995-37140/c  
;/ Sequence 37140, Application US/09918995  
;/ Publication No. US20030073623A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Eyseq, Inc.  
;/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
;/ FILE REFERENCE: 20411-756

;/ CURRENT APPLICATION NUMBER: US/09/918,995  
;/ CURRENT FILING DATE: 2001-07-30  
;/ PRIOR APPLICATION NUMBER: US/09/235,076  
;/ PRIOR FILING DATE: 1999-01-20  
;/ NUMBER OF SEQ ID NOS: 38054  
;/ SOFTWARE: FastSeq for Windows Version 3.0  
;/ SEQ ID NO 37140  
;/ LENGTH: 490  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: misc feature  
;/ LOCATION: (1)-(490)  
;/ OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-37140

Query Match 100.0%; Score 21; DB 10; Length 490;  
Best Local Similarity 100.0%; Pred. No. 0.69;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||  
Db 277 CCCTGTTGGTCTATGCGCTG 257

RESULT 37  
US-10-102-524-715/c  
;/ Sequence 715, Application US/10102524  
;/ Publication No. US20030109434A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Algate, Paul A.  
;/ APPLICANT: Mannion, Jane  
;/ APPLICANT: Gaiger, Alexander  
;/ APPLICANT: Gordon, Brian  
;/ APPLICANT: Harlocker, Susan L.  
;/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
;/ THERAPY AND DIAGNOSIS OF KIDNEY CANCER  
;/ FILE REFERENCE: 210121.572  
;/ CURRENT APPLICATION NUMBER: US/10/102,524  
;/ CURRENT FILING DATE: 2002-03-19  
;/ NUMBER OF SEQ ID NOS: 1863  
;/ SOFTWARE: FastSeq for Windows Version 4.0  
;/ SEQ ID NO 715  
;/ LENGTH: 523  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
US-10-102-524-715

Query Match 100.0%; Score 21; DB 15; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.69;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGCTG 21  
|||  
Db 192 CCCTGTTGGTCTATGCGCTG 172

RESULT 38  
US-10-084-817-88/c  
;/ Sequence 88, Application US/10084817  
;/ Publication No. US20030119009A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Susan Stuart  
;/ APPLICANT: Jed G. Nuchtern  
;/ APPLICANT: Sharon E. Pion  
;/ APPLICANT: Jason M. Shohet  
;/ TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
;/ FILE REFERENCE: PA-0046 US  
;/ CURRENT APPLICATION NUMBER: US/10/084,817  
;/ CURRENT FILING DATE: 2002-02-25  
;/ PRIOR APPLICATION NUMBER: 60/270,784  
;/ PRIOR FILING DATE: 2001-02-23  
;/ NUMBER OF SEQ ID NOS: 365

; SOFTWARE: PERL Program  
; SEQ ID NO 88  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US2003011909A1 1090035.1  
US-10-084-817-88

Query Match 100.0%; Score 21; DB 15; Length 601;  
Best Local Similarity 100.0%; Pred. No. 0.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
Db 225 CCCTGTTGGTCTATGCGTCTG 205

## RESULT 39

US-10-220-120-138/c  
; Sequence 138, Application US/10220120  
; Publication No. US20040048253A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: PANZER, Scott R.  
; APPLICANT: SPIRO, Peter A.  
; APPLICANT: BANVILLE, Steven C.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: CHEALUP, Michael S.  
; APPLICANT: CHANG, Simon C.  
; APPLICANT: CHEN, Alice  
; APPLICANT: D'SA, Steven A.  
; APPLICANT: AMSHEY, Stefan  
; APPLICANT: DAHL, Christopher R.  
; APPLICANT: DAM, Tam C.  
; APPLICANT: DANIELS, Susan E.  
; APPLICANT: DUFOUR, Gerard E.  
; APPLICANT: FLORES, Vincent  
; APPLICANT: FONG, Willy T.  
; APPLICANT: GREENAWALT, Lila B.  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: JONES, Anissa L.  
; APPLICANT: LIU, Tommy F.  
; APPLICANT: ROSEBERRY, Ann M.  
; APPLICANT: ROSEN, Bruce H.  
; APPLICANT: RUSSO, Frank D.  
; APPLICANT: STOCKREHER, Theresa K.  
; APPLICANT: DAPFO, Abel  
; APPLICANT: WRIGHT, Rachel J.  
; APPLICANT: YAP, Pierre E.  
; APPLICANT: YU, Jimmy Y.  
; APPLICANT: BRADLEY, Diana L.  
; APPLICANT: BRATCHER, Shawn R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: COHEN, Howard J.  
; APPLICANT: HODGSON, David M.  
; APPLICANT: LINCOLN, Stephen E.  
; APPLICANT: JACKSON, Stuart  
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: FT-1113 PCT  
; CURRENT APPLICATION NUMBER: US/10/220,120  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,773; 60/184,776;  
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;  
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;  
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;  
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;  
; 60/205,324; 60/205,286  
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;

; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;  
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
; 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;  
; 2000-05-17; 2000-05-17  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: PERL Program  
; SEQ ID NO 138  
; LENGTH: 1259  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040048253A1 Li:1169865.1:2000MAY01  
US-10-220-120-138

Query Match 100.0%; Score 21; DB 13; Length 1259;  
Best Local Similarity 100.0%; Pred. No. 0.74;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21  
Db 367 CCCTGTTGGTCTATGCGTCTG 347

## RESULT 40

US-09-925-302-351/c  
; Sequence 351, Application US/09925302  
; Patent No. US20020044941A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSEN et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 351  
; LENGTH: 1348  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (12)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (24)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (36)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1294)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1307)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1318)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1329)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-302-351

Query Match 100.0%; Score 21; DB 9; Length 1348;  
Best Local Similarity 100.0%; Pred. No. 0.74;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATGCGTCTG 21

```
Db 305 CCCTGTTGGTCTATCGGTCGTG 285
|||||
RESULT 41
US-09-925-302-351/c
; Sequence 351, Application US/09925302
; Publication No. US20030364072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1294)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1318)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1329)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-302-351
Query Match 100.0%; Score 21; DB 13; Length 1348;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGTCGTG 21
|||||
Db 305 CCCTGTTGGTCTATCGGTCGTG 285
|||||
RESULT 42
US-09-960-706-1043/c
; Sequence 1043, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 691
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
; US-09-873-319-691
Query Match 100.0%; Score 21; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGTCGTG 21
|||||
Db 6386 CCCTGTTGGTCTATCGGTCGTG 6366
|||||
RESULT 43
US-09-873-319-691/c
; Sequence 691, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 691
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
; US-09-873-319-691
Query Match 100.0%; Score 21; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTTGGTCTATCGGTCGTG 21
|||||
Db 6386 CCCTGTTGGTCTATCGGTCGTG 6366
|||||
RESULT 44
US-09-877-819B-39/c
; Sequence 39, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: White, Scott
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,864
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
; US-09-960-706-1043
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Primer sequence
US-09-877-819B-39

Query Match          95.2%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCTGTGGTCTATGGCTCTG 21
Db      20 CCTGTGGTCTATGGCTCTG 1

RESULT 45
US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 Y61855
US-09-917-800A-1424
```

```
Query Match          90.5%; Score 19; DB 9; Length 4254;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCTGTGGTCTATGGCTCT 20
Db      1008 CCTGTGGTCTATGGCTCT 1026
```

```
RESULT 46
US-10-388-934-265
; Sequence 265, Application US/10388934
```

```
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Soess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 265
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-265
```

```
Query Match          90.5%; Score 19; DB 16; Length 4254;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCTGTGGTCTATGGCTCT 20
Db      1008 CCTGTGGTCTATGGCTCT 1026
```

```
RESULT 47
US-09-894-844-65
; Sequence 65, Application US/09894844
; Patent No. US20020176873A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAM102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-65
```

```
Query Match          84.8%; Score 17.8; DB 9; Length 1860;
Best Local Similarity 90.5%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCTGTGGTCTATGGCTCTG 21
Db      1255 CCTGTGGTCTATGGCTCTG 1275
```

```
RESULT 48
US-10-647-089-65
; Sequence 65, Application US/10647089
; Publication No. US20040063923A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
```



```
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: the M. Tuberculosis Complex
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1998-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-10-647-089-65

Query Match      84.8%; Score 17.8; DB 13; Length 1860;
Best Local Similarity 90.5%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATGCGTCTG 21
DB 1255 CCTGTGCTCTATGCGTCTG 1275

RESULT 49
US-10-388-902-65
; Sequence 65, Application US/10388902
; Publication No. US20040018574A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAN102CON
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1998-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-10-388-902-65

Query Match      84.8%; Score 17.8; DB 16; Length 1860;
Best Local Similarity 90.5%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGCTCTATGCGTCTG 21
DB 1255 CCTGTGCTCTATGCGTCTG 1275

RESULT 50
US-10-647-089-65
; Sequence 65, Application US/10388902
; Publication No. US20040018574A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAN102CON
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1998-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-10-647-089-65

Query Match      84.8%; Score 17.8; DB 13; Length 1860;
Best Local Similarity 90.5%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46603
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46603

Query Match      82.9%; Score 17.4; DB 13; Length 696;
Best Local Similarity 94.7%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CTGTGGTCTATGCGTCTG 21
DB 218 CTGTGGTCTATGCTCTG 236

RESULT 51
US-10-027-632-46603
; Sequence 46603, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46603
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46603

Query Match      82.9%; Score 17.4; DB 16; Length 696;
Best Local Similarity 94.7%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CTGTGGTCTATGCGTCTG 21
```

```
Db      218  CTGTTGGTCTATGTCGTCG 236
|||||
RESULT 52
US-10-027-632-222275
; Sequence 222275, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222275
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222275

Query Match      80.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  CCTGTTGGTCTATGCGTCGTCG 21
||| |||||
Db      314  CCTTTGGTCTATGTCGTCG 333

RESULT 53
US-10-027-632-222276
; Sequence 222276, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222276
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222276

Query Match      80.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  CCTGTTGGTCTATGCGTCGTCG 21
||| |||||
Db      314  CCTTTGGTCTATGTCGTCG 333
```

```
; SEQ ID NO 222276
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222276

Query Match      80.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  CCTGTTGGTCTATGCGTCGTCG 21
||| |||||
Db      314  CCTTTGGTCTATGTCGTCG 333

RESULT 54
US-10-027-632-222277
; Sequence 222277, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222277
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222277

Query Match      80.0%; Score 16.8; DB 13; Length 572;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  CCTGTTGGTCTATGCGTCGTCG 21
||| |||||
Db      314  CCTTTGGTCTATGTCGTCG 333

RESULT 55
US-10-027-632-222278
; Sequence 222278, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222278  
; LENGTH: 572  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-222278

Query Match 80.0%; Score 16.8; DB 13; Length 572;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTGTTGGTCTATGCGTCTG 21  
||| ||||| ||||| |||||  
Db 314 CCTTTGGTCTATGTCGTG 333

RESULT 56  
US-10-027-632-222275  
; Sequence 222275, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222275  
; LENGTH: 572  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-222275

Query Match 80.0%; Score 16.8; DB 16; Length 572;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTGTTGGTCTATGCGTCTG 21  
||| ||||| ||||| |||||  
Db 314 CCTTTGGTCTATGTCGTG 333

RESULT 57  
US-10-027-632-222276  
; Sequence 222276, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222276  
; LENGTH: 572  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-222276

Query Match 80.0%; Score 16.8; DB 16; Length 572;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTGTTGGTCTATGCGTCTG 21  
||| ||||| ||||| |||||  
Db 314 CCTTTGGTCTATGTCGTG 333

RESULT 58  
US-10-027-632-222277  
; Sequence 222277, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222277  
; LENGTH: 572  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-222277

Query Match 80.0%; Score 16.8; DB 16; Length 572;  
Best Local Similarity 90.0%; Pred. No. 99;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 604.235 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819B-33

Perfect score: 20

Sequence: 1 gagcagccctctctcttct 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_man.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	18.4	92.0	139056	2	AC141691
C 2	18.4	92.0	188090	9	AC011625
C 3	18.4	92.0	219368	10	AC098721
C 4	18	90.0	133984	9	AL512306
C 5	18	90.0	153033	2	AC021462
C 6	17.4	37.0	29134	3	CER186
C 7	17	35.0	124479	2	AC020966
C 8	16.8	84.0	1014	1	TY14655
C 9	16.8	84.0	3312	5	CHKPLB
C 10	16.8	84.0	3423	5	CHKPLB2
C 11	16.8	84.0	3421	8	AK05803
C 12	16.8	84.0	6176	9	F247811S12
C 13	16.8	84.0	28396	3	U41264
C 14	16.8	84.0	33010	5	AL592204
C 15	16.8	84.0	52423	2	AC020395
C 16	16.8	84.0	71503	9	AF219991
C 17	16.8	84.0	80423	3	DROABDB
C 18	16.8	84.0	118948	2	AP003226
C 19	16.8	84.0	121088	8	AP004326
C 20	16.8	84.0	125364	4	AC139728
C 21	16.8	84.0	132060	2	AC138535
C 22	16.8	84.0	132900	8	AC113948
C 23	16.8	84.0	136548	8	AP004368
C 24	16.8	84.0	142472	9	AC142304
C 25	16.8	84.0	144595	8	AP004574
C 26	16.8	84.0	148012	2	AC138995
C 27	16.8	84.0	149061	9	AC098805
C 28	16.8	84.0	151071	8	AP003449
C 29	16.8	84.0	151249	2	AC135250
C 30	16.8	84.0	151334	2	EX571897
C 31	16.8	84.0	152316	10	AC007844
C 32	16.8	84.0	152802	8	AP005491
C 33	16.8	84.0	156173	2	AC109792
C 34	16.8	84.0	157358	9	AC009163
C 35	16.8	84.0	160613	2	AC134969
C 36	16.8	84.0	165165	3	AC095018
C 37	16.8	84.0	175335	3	AC091636
C 38	16.8	84.0	176552	9	AC106820
C 39	16.8	84.0	179886	10	AC114826
C 40	16.8	84.0	183037	2	AC136111
C 41	16.8	84.0	183538	2	AC147272
C 42	16.8	84.0	194361	8	AC092389
C 43	16.8	84.0	194832	9	AC025287
C 44	16.8	84.0	201853	9	AL591424
C 45	16.8	84.0	203284	2	AC110130
C 46	16.8	84.0	205268	2	AC146518
C 47	16.8	84.0	212221	2	AC136126
C 48	16.8	84.0	227877	2	AC109725
C 49	16.8	84.0	228544	3	AE003715
C 50	16.8	84.0	230161	2	AC128213
C 51	16.8	84.0	232945	2	AC103156
C 52	16.8	84.0	238596	2	AC119587
C 53	16.8	84.0	239080	2	AC111635
C 54	16.8	84.0	240550	2	AC111385
C 55	16.8	84.0	241178	2	AC130508
C 56	16.8	84.0	244843	2	AC107162
C 57	16.8	84.0	252978	2	AC126983
C 58	16.8	84.0	266661	2	AC126523
C 59	16.8	84.0	295712	2	AC120456
C 60	16.8	84.0	296282	2	AC111857
C 61	16.8	84.0	300029	8	AE017114
C 62	16.8	84.0	316828	2	AC135386
C 63	16.8	84.0	338234	3	DM031961
C 64	16.4	82.0	37068	8	AC067938
C 65	16.4	82.0	43556	8	AC067937

C 66 16.4 82.0 59429 2 AC091176 Homo sapi  
 C 67 16.4 82.0 63953 2 AC087663 Homo sapi  
 C 68 16.4 82.0 68217 9 AL590664 Human DNA  
 C 69 16.4 82.0 69208 2 AC020466 Drosophila  
 C 70 16.4 82.0 74512 2 AC100521 Mus muscu  
 C 71 16.4 82.0 73593 8 AF263283 Filobasid  
 C 72 16.4 82.0 93979 8 AC068564 Filobasid  
 C 73 16.4 82.0 95663 9 AC010247 Homo sapi  
 C 74 16.4 82.0 126323 8 AC132215 Genomic s  
 C 75 16.4 82.0 159320 2 AUL38757 Homo sapi  
 C 76 16.4 82.0 167108 2 AC068283 Homo sapi  
 C 77 16.4 82.0 175118 3 AC010842 Drosophila  
 C 78 16.4 82.0 205488 2 AC115975 Mus muscu  
 C 79 16.4 82.0 220371 2 AC109966 Rattus no  
 C 80 16.4 82.0 240051 2 AC123236 Mus muscu  
 C 81 16.4 82.0 268335 2 AC126066 Rattus no  
 C 82 16.4 82.0 295225 3 AB003461 Drosophila  
 C 83 16.4 82.0 303862 1 AB017215 Geobacter  
 C 84 16.4 80.0 155164 9 AC093182 Homo sapi  
 C 85 16.4 80.0 176209 9 AC009518 Homo sapi  
 C 86 16.4 80.0 195811 2 AC147038 Pan trogl  
 C 87 16.4 80.0 213604 2 AC100345 Mus muscu  
 C 88 16.4 80.0 236527 2 AC147103 Pan trogl  
 C 89 16.4 80.0 274669 2 BX005328 Drosophila  
 C 90 15.8 79.0 668 8 CNS01944 Homo sapi  
 C 91 15.8 79.0 720 8 CNS01944 Homo sapi  
 C 92 15.8 79.0 780 8 CNS01944 Homo sapi  
 C 93 15.8 79.0 780 8 CNS01944 Homo sapi  
 C 94 15.8 79.0 869 3 SPH2B1G Homo sapi  
 C 95 15.8 79.0 951 8 AK103638 Oryza sat  
 C 96 15.8 79.0 1092 9 HSA403950 Homo sapi  
 C 97 15.8 79.0 1362 6 AX431819 Sequence  
 C 98 15.8 79.0 1770 10 BC013462 Homo sapi  
 C 99 15.8 79.0 2000 6 AX656405 Sequence  
 C 100 15.8 79.0 2000 6 AX656673 Sequence  
 C 101 15.8 79.0 3059 8 AK069509 Oryza sat  
 C 102 15.8 79.0 32-3 9 HSEXMONTOR Homo sapi  
 C 103 15.8 79.0 3770 8 AK121357 Oryza sat  
 C 104 15.8 79.0 3882 10 BC057380 Mus muscu  
 C 105 15.8 79.0 6974 3 NSPRTRANS Homo sapi  
 C 106 15.8 79.0 7690 6 AX346024 Sequence  
 C 107 15.8 79.0 7950 1 AF016298 Rhodospirillum rubrum  
 C 108 15.8 79.0 11002 1 AF117827 Methylobacterium thermophilum  
 C 109 15.8 79.0 14924 6 AX281280 Sequence  
 C 110 15.8 79.0 14924 6 AX281280 Sequence  
 C 111 15.8 79.0 38951 3 AC105441 Leishmania  
 C 112 15.8 79.0 39949 3 AC009782 Leishmania  
 C 113 15.8 79.0 40679 9 AC002052 Leishmania  
 C 114 15.8 79.0 58020 2 AF002772 Homo sapi  
 C 115 15.8 79.0 58852 2 AC036190 Homo sapi  
 C 116 15.8 79.0 61076 2 AC104860 Mus muscu  
 C 117 15.8 79.0 64401 2 AC131908 Homo sapi  
 C 118 15.8 79.0 64523 2 AC100987 Homo sapi  
 C 119 15.8 79.0 57427 2 AC101742 Homo sapi  
 C 120 15.8 79.0 70752 2 AC113148 Homo sapi  
 C 121 15.8 79.0 7-241 2 AC102061 Mus muscu  
 C 122 15.8 79.0 71241 2 AC102061 Mus muscu  
 C 123 15.8 79.0 72400 8 AP005524 Oryza sat  
 C 124 15.8 79.0 78770 2 AC034230 Homo sapi  
 C 125 15.8 79.0 80465 2 AC034230 Homo sapi  
 C 126 15.8 79.0 82270 9 AL354879 Human DNA  
 C 127 15.8 79.0 86722 9 AB012245 Arabidopsis  
 C 128 15.8 79.0 91826 9 AB045320 Homo sapi  
 C 129 15.8 79.0 92455 2 AC034136 Homo sapi  
 C 130 15.8 79.0 92510 9 HS39083 Homo sapi  
 C 131 15.8 79.0 93240 2 AC007586 Homo sapi  
 C 132 15.8 79.0 95597 2 AC014022 Homo sapi  
 C 133 15.8 79.0 106256 3 AC108135 Homo sapi  
 C 134 15.8 79.0 108533 8 AF527809 Sorghum b  
 C 135 15.8 79.0 108553 8 AF527809 Sorghum b  
 C 136 15.8 79.0 110000 2 AC112799 Rattus no  
 C 137 15.8 79.0 110000 2 AC112799 Rattus no  
 C 138 15.8 79.0 110000 2 LMFLCHR36\_28 Continuation (9 of

C 139 15.8 79.0 110000 2 LMFLCHR36\_28 Continuation (29 of  
 C 140 15.8 79.0 111111 10 AC130816 Mus muscu  
 C 141 15.8 79.0 112630 2 AC025461 Homo sapi  
 C 142 15.8 79.0 119171 2 BX247904 Drosophila  
 C 143 15.8 79.0 121041 9 AC008042 Homo sapi  
 C 144 15.8 79.0 121041 9 AC008042 Homo sapi  
 C 145 15.8 79.0 122940 8 AP004054 Oryza sat  
 C 146 15.8 79.0 123013 2 AC010005 Drosophila  
 C 147 15.8 79.0 127218 2 AC141738 Apis mell  
 C 148 15.8 79.0 127675 9 AC146221 Pan trogl  
 C 149 15.8 79.0 130027 8 AC009773 Homo sapi  
 C 150 15.8 79.0 130193 8 AC138004 Oryza sat  
 C 151 15.8 79.0 132910 2 AC014319 Drosophila  
 C 152 15.8 79.0 133713 3 AC087838 Homo sapi  
 C 153 15.8 79.0 135940 2 AC119406 Trypanosoma  
 C 154 15.8 79.0 140100 9 AC016778 Homo sapi  
 C 155 15.8 79.0 141489 2 AC141017 Rattus no  
 C 156 15.8 79.0 142195 2 AC010669 Drosophila  
 C 157 15.8 79.0 146158 2 AF004868 Oryza sat  
 C 158 15.8 79.0 146717 3 AC097726 Drosophila  
 C 159 15.8 79.0 149041 9 AC044782 Homo sapi  
 C 160 15.8 79.0 149716 2 AC119380 Rattus no  
 C 161 15.8 79.0 150377 2 AC093405 Homo sapi  
 C 162 15.8 79.0 150995 8 AC104615 Oryza sat  
 C 163 15.8 79.0 151967 10 AL805912 Mouse DNA  
 C 164 15.8 79.0 153149 9 AL354819 Human DNA  
 C 165 15.8 79.0 153508 2 AC141790 Apis mell  
 C 166 15.8 79.0 154954 9 AC022507 Homo sapi  
 C 167 15.8 79.0 156221 2 AC134963 Canis fam  
 C 168 15.8 79.0 159516 2 AC048350 Homo sapi  
 C 169 15.8 79.0 160252 2 AC1390250 Homo sapi  
 C 170 15.8 79.0 160785 9 AC065583 Homo sapi  
 C 171 15.8 79.0 161741 8 AC120535 Oryza sat  
 C 172 15.8 79.0 161852 2 AC135690 Rattus no  
 C 173 15.8 79.0 162083 2 AC124654 Homo sapi  
 C 174 15.8 79.0 163043 9 AL591069 Human DNA  
 C 175 15.8 79.0 163066 2 AC116506 Mus muscu  
 C 176 15.8 79.0 163475 10 AC026761 Mus muscu  
 C 177 15.8 79.0 164812 9 AC078820 Homo sapi  
 C 178 15.8 79.0 165594 2 AC023317 Homo sapi  
 C 179 15.8 79.0 167408 2 AC025526 Homo sapi  
 C 180 15.8 79.0 167711 2 AC079166 Homo sapi  
 C 181 15.8 79.0 168990 2 AC122516 Mus muscu

## ALIGNMENTS

## RESULT 1

AC141691/c

LOCUS

DEFINITION

AC141691

APIS mellifera clone CH224-5515, WORKING DRAFT SEQUENCE, 30

unordered pieces.

AC141691

VERSION

AC141691.1 GI:29123875

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Apis mellifera (honeybee)

ORGANISM

Apis mellifera

Bukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Hymenoptera; Apoidea; Apidae;

Apidae; Apis.

REFERENCE

1 (bases 1 to 139056)

AUTHORS

Muzny, D.M., Adams, C., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Anarlunga, H.C., Are, J.R., Ayale, M., Banks, T.,

Barbieri, J., Benton, J., Binage, K., Blankenburg, K., Bonin, D.,

Bouck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Carroll, T., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

DeVall, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Continuation (9 of

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsai, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, U., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, M., Loughheed, H., Lozano, R. J., Lu, X., Lucher, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, X., Morris, S., Moser, M., Neai, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, E., Sutton, A., Svatek, A., Tabrizi, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, O., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmari, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 139056)  
Worley, K.C.

Submitted (19-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: AMGP  
Center clone name: CH224-5315  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 125289 bases at least Q40  
Consensus quality: 128700 bases at least Q30  
Consensus quality: 130487 bases at least Q20  
Estimated insert size: 129688; sum-of-contigs estimation  
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

\*\*\*\*\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\*\*\*\*\* NOTE: This is a 'working draft' sequence. It currently consists of 30 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\*\*\*\*\*  
1 1555: contig of 1555 bp in length  
\* 1556  
\* 1555: gap of unknown length  
\* 1656  
\* 3079: contig of 1424 bp in length  
\* 3080  
\* 3179: gap of unknown length  
\* 3180  
\* 4896: contig of 1717 bp in length  
\* 4897  
\* 4936: gap of unknown length  
\* 4897  
\* 6870: contig of 1874 bp in length  
\* 6871  
\* 6970: gap of unknown length  
\* 6971  
\* 8284: contig of 1314 bp in length

8285 8384: gap of unknown length  
\* 8385  
\* 10898: contig of 2514 bp in length  
\* 10899  
\* 12432: contig of 1434 bp in length  
\* 12433  
\* 14591: contig of 2059 bp in length  
\* 14592  
\* 14591: gap of unknown length  
\* 14592  
\* 16015: contig of 1324 bp in length  
\* 16016  
\* 16115: gap of unknown length  
\* 16116  
\* 18336: contig of 2721 bp in length  
\* 18337  
\* 20628: contig of 1692 bp in length  
\* 20629  
\* 23613: contig of 2895 bp in length  
\* 23614  
\* 23713: gap of unknown length  
\* 23714  
\* 26592: contig of 2879 bp in length  
\* 26593  
\* 29806: contig of 3114 bp in length  
\* 29807  
\* 29906: gap of unknown length  
\* 29907  
\* 34149: contig of 4243 bp in length  
\* 34150  
\* 34249: gap of unknown length  
\* 34250  
\* 37904: contig of 3655 bp in length  
\* 37905  
\* 38004: gap of unknown length  
\* 38005  
\* 4018: contig of 2914 bp in length  
\* 4019  
\* 4018: gap of unknown length  
\* 4019  
\* 45267: contig of 4249 bp in length  
\* 45268  
\* 45367: gap of unknown length  
\* 45368  
\* 50380: contig of 5213 bp in length  
\* 50381  
\* 55750: contig of 5070 bp in length  
\* 55751  
\* 55850: gap of unknown length  
\* 55851  
\* 60417: contig of 4567 bp in length  
\* 60418  
\* 60517: gap of unknown length  
\* 60518  
\* 65870: contig of 5353 bp in length  
\* 65871  
\* 70676: contig of 4706 bp in length  
\* 70677  
\* 70776: gap of unknown length  
\* 70777  
\* 77788: contig of 7012 bp in length  
\* 77789  
\* 77888: gap of unknown length  
\* 77889  
\* 82008: contig of 4120 bp in length  
\* 82009  
\* 82108: gap of unknown length  
\* 82109  
\* 88870: contig of 6762 bp in length  
\* 88871  
\* 88970: gap of unknown length  
\* 88971  
\* 98432: contig of 9462 bp in length  
\* 98433  
\* 98532: gap of unknown length  
\* 98533  
\* 108197  
\* 108296: gap of unknown length  
\* 108297  
\* 122360: contig of 14064 bp in length  
\* 122361  
\* 122461  
\* 139056: contig of 16596 bp in length.

Location/Qualifiers  
1.139056  
/organism="Apis mellifera"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7460"  
/clone="CH224-5515"

Query Match 92.0%; Score 18.4; DB 2; Length 139056;  
Best Local Similarity 95.0%; Pred. No. 73;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACGACGGCTCTCTTCTTCT 20  
|||||  
Db 91730 GACGACGGCTCTCTTCTTCT 91711  
|||||

RESULT 2  
AC011625  
LOCUS AC011625 188090 bp DNA linear PRI 08-DEC-1999  
DEFINITION Homo sapiens clone RP11-341C17 from 7q31, complete sequence.  
AC011625  
AC011625.2 GI:6539285

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

KEYWORDS  
SOURCE  
ORGANISM

HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 188090)  
Kaul,R.K., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P.  
and Olson,M.V.  
Large-scale MCD Mapping and Sequencing of Human Chromosome 7  
Unpublished  
2 (bases 1 to 188090)  
Bubb,K.L., Desmarais,C.L., Ramsey,S.A. and Hubley,R.M.  
Direct Submission  
Submitted (08-OCT-1999) Human Genome Center, University of  
Washington, Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 188090)  
Kaul,R.K. and Desmarais,C.L.  
Direct Submission  
Submitted (08-DEC-1999) Human Genome Center, University of  
Washington, Box 352145, Seattle, WA 98195, USA  
On Dec 8, 1999 this sequence version replaced gi:6016750.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

----- Genome Center:  
University of Washington Genome Center  
Center code: UWGC  
Web site: <http://genome.washington.edu>  
Contact: [uwgchelp@u.washington.edu](mailto:uwgchelp@u.washington.edu)  
----- Project Information  
Center project name: chr-7  
Center clone name: djs301 (RP11-341C17)  
----- Summary Statistics

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Sequencing vector: M13; 100% of reads  
Chemistry: Dye-Primer Bodypy; 93% of reads Chemistry:  
Dye-terminator Big Dye; 7% of reads  
Assembly program: Phrap; version 0.990319  
Insert size: 188 094; sum-of-contigs  
Quality coverage: 7.78x in Q20 bases; sum-of-contigs

## Overlapping Sequences:

5': UWGC:djs302  
3': UWGC:djs380

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

## Double stranded (DS) coverage:

DS or two chemistry coverage: 70.4%  
Single stranded regions: 100.0%

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-600 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

FP	BgIII	Seq	FP	Seq	FP	Seq
2095.00	2061.00	8947.00	8687.00	729.00	732.00	
12518.00	12347.00	844.00	812.00	1463.00	1480.00	

5082.00	5007.00	4657.00	4593.00	5528.00	5438.00
2694.00	2672.00	3637.00	3613.00	3383.00	3371.00
5175.00	5083.00	1152.00	1147.00	1240.00	1220.00
1094.00	1093.00	11517.00	11411.00	4875.00	4844.00
2231.00	2189.00	16729.00	16537.00	4194.00	4094.00
16006.00	15888.00	13550.00	13383.00	2402.00	2353.00
6982.00	6668.00	5001.00	4900.00	1463.00	1439.00
8822.00	8629.00	6353.00	6107.00	2072.00	2056.00
11945.00	11811.00	3101.00	3027.00	5301.00	5143.00
10740.00	10587.00	3850.00	3719.00	1463.00	1408.00
2095.00	2062.00	1925.00	1935.00	729.00	706.00
7981.00	7820.00	6169.00	6045.00	10419.00	10239.00
2095.00	2101.00	6353.00	6206.00	4194.00	4107.00
1178.00	1141.00	1743.00	1714.00	1128.00	1118.00
3623.00	3525.00	2067.00	2052.00	622.00	630.00
621.00	581.00	559.00	569.00	1045.00	1052.00
6428.00	6341.00	544.00	510.00	7407.00	7153.00
6148.00	6023.00	2364.00	2354.00	3498.00	3389.00
769.00	736.00	660.00	633.00	4506.00	4428.00
4722.00	4640.00	1013.00	972.00	2402.00	2339.00
1178.00	1180.00	11962.00	11788.00	13094.00	12940.00
7083.00	6917.00	5201.00	5141.00	4378.00	4306.00
2095.00	2022.00	5201.00	5091.00	2025.00	1988.00
2429.00	2372.00	1925.00	1902.00	4037.00	3911.00
2095.00	2083.00	1925.00	1853.00	2402.00	2326.00
546.00	544.00	2967.00	2843.00	2072.00	2069.00
4240.00	4182.00	3850.00	3715.00	1609.00	1583.00
3913.00	3829.00	660.00	652.00	2922.00	2845.00
1656.00	1641.00	29666.00	26384.00	5301.00	5252.00
621.00	623.00	544.00	531.00	1747.00	1719.00
1940.00	1927.00	1551.00	1536.00	729.00	697.00
18696.00	18261.00	1152.00	1097.00	2646.00	2580.00
3153.00	3085.00	1925.00	1922.00	2402.00	2335.00
2966.00	2898.00	14113.00	14024.00	10419.00	10108.00
2429.00	2419.00	6876.00	6794.00	7407.00	7317.00
3913.00	3763.00	1925.00	1869.00	8785.00	8640.00



1524.00 1500.00 6169.00 6048.00 7052.00 6870.00  
 3776.00 3636.00 2207.00 2144.00 3731.00 3659.00  
 5704.00 5638.00 4194.00 4085.00  
 6982.00 6868.00 622.00 581.00  
 3881.00 3751.00  
 1240.00 1221.00  
 1319.00 1284.00  
 5528.00 5416.00  
 7052.00 6844.00  
 3881.00 3888.00  
 3498.00 3426.00  
 14516.00 14342.00

## FEATURES

source

Location/Qualifiers  
 1. 188090  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7q31"  
 /clone="RP11-341C17 or UWGC:djs301"  
 /cell\_line="Male Blood"  
 /clone\_lib="RPC-11 Human Male BAC Library"  
 444.750  
 /rpt\_family="Alu"  
 891.963  
 /standard\_name="SWS92952"  
 /notes="GenBank Accession: G13043"  
 complement(2131..2409)  
 /rpt\_family="Alu"  
 complement(8590..8831)  
 /rpt\_family="Alu"  
 10548..10834  
 /rpt\_family="Alu"  
 11525..11795  
 /rpt\_family="Alu"  
 18103..18403  
 /rpt\_family="Alu"  
 complement(27089..27369)  
 Query Match 92.08; Score 18.4; DB 9; Length 185090;  
 Best Local Similarity 95.08; Pred.No. 73;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACGACCCCTCTCTTCT 20

Db 80765 GACGACCCCTCTCTTCT 80784

## RESULT 3

LOCUS

AC098721

DEFINITION

Mus musculus BAC clone RP23-2023 from 2, complete sequence.

ACCESSION

AC098721.3

VERSION

GI:19909472

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

REFERENCE

1 (bases 1 to 219368)

AUTHORS

Armstrong,J., Abbott,A. and Creason,K.

TITLE  
JOURNAL  
REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The sequence of Mus musculus BAC clone RP23-2023

Unpublished (2001)

2 (bases 1 to 219368)

Wilson,R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 219368)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (31-OCT-2001) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 219368)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (03-APR-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 219368)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 219368)

Wilson,R.

Direct Submission

Submitted (31-OCT-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Apr 3, 2002 this sequence version replaced gi:17017625.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

----- Summary Statistics

-----

Center project name: M\_PA0002023

-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

## SOURCE INFORMATION:

The RP23-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Tateno in the laboratory of Pieter de Jong (http://www.chori.org) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

1. 219368

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="2"

/map="2"

/clone="RP23-2023"

## FEATURES

source

```

repeat_region /clone_lib="RPCT-23"
repeat_region 4..99
repeat_region /rpt_family="Alu"
repeat_region 1931..2204
repeat_region /rpt_family="L2"
repeat_region 3672..3749
repeat_region /rpt_family="Alu"
repeat_region 3832..3971
repeat_region /rpt_family="Alu"
repeat_region 6380..6576
repeat_region /rpt_family="B2"
repeat_region 7101..7291
repeat_region /rpt_family="B2"
repeat_region 9747..9855
repeat_region /rpt_family="L1"
repeat_region 9914..10361
repeat_region /rpt_family="L1"
repeat_region 17647..17740
repeat_region /rpt_family="MIR"
repeat_region 19477..19723
repeat_region /rpt_family="B4"
repeat_region 19721..19824
repeat_region /rpt_family="B4"
repeat_region 20827..21403
repeat_region /rpt_family="L1"
repeat_region 21442..22029
repeat_region /rpt_family="L1"
misc_feature 23389..23433
repeat_region /note="Unresolved homopolymeric repeat."
repeat_region 23652..23763
repeat_region /rpt_family="L1"
repeat_region 23791..23925
repeat_region /rpt_family="Alu"
repeat_region 24038..24740
repeat_region /rpt_family="L1"
repeat_region 24731..25012
repeat_region /rpt_family="L1"
repeat_region 25372..25932
repeat_region /rpt_family="L1"
repeat_region 25926..26093
repeat_region /rpt_family="L1"
repeat_region 26350..26745
repeat_region /rpt_family="MaLR"
repeat_region 26896..27019
repeat_region /rpt_family="ERVK"
repeat_region 26993..27010
repeat_region /rpt_family="B4"
repeat_region 27011..27095
repeat_region /rpt_family="Alu"
repeat_region 27128..27495
repeat_region /rpt_family="ERVK"
repeat_region 27842..28110
repeat_region /rpt_family="MaLR"
repeat_region 28214..28446
repeat_region /rpt_family="L1"
repeat_region 30377..30522
repeat_region /rpt_family="Alu"
repeat_region 31240..31636
repeat_region /rpt_family="L1"
repeat_region 31631..31972
repeat_region /rpt_family="L1"
repeat_region 32577..32784
repeat_region /rpt_family="B2"
repeat_region 33405..33511
repeat_region /rpt_family="L1"
repeat_region 34749..35091
repeat_region /rpt_family="L1"
repeat_region 36138..36280
repeat_region /rpt_family="ERVK"
repeat_region 37095..37481
repeat_region /rpt_family="L2"
repeat_region 37875..37968
repeat_region /rpt_family="MIR"

```

```

repeat_region 44056..44148
repeat_region /rpt_family="Alu"
repeat_region 45143..45273
repeat_region /rpt_family="MER1_type"
repeat_region 45296..45504
repeat_region /rpt_family="MaLR"
repeat_region 45546..45954
repeat_region /rpt_family="L1"
repeat_region 45982..47257
repeat_region /rpt_family="MER1_type"
repeat_region 47306..47380
repeat_region /rpt_family="tRNA-Ala-GCY_"
repeat_region 47394..47511
repeat_region /rpt_family="MER1_type"
repeat_region 49498..49991
repeat_region /rpt_family="ERVK"
repeat_region 50037..50184
repeat_region /rpt_family="Alu"
repeat_region 52307..53184
repeat_region /rpt_family="L1"
repeat_region 53851..53837
repeat_region /rpt_family="L1"
repeat_region 53848..53958
repeat_region /rpt_family="5S"
repeat_region 53985..54117
repeat_region /rpt_family="L1"
repeat_region 54119..54567
repeat_region /rpt_family="RMER6B"
repeat_region 55344..55466
repeat_region /rpt_family="ERVK"
repeat_region 55852..56355
repeat_region /rpt_family="L1"
repeat_region 56390..56440
repeat_region /rpt_family="Alu"
repeat_region 58409..59200
repeat_region /rpt_family="ERVK"
repeat_region 59202..60566
repeat_region /rpt_family="L1"
repeat_region 60573..60633
repeat_region /rpt_family="ERV1"
repeat_region 60753..61495
repeat_region /rpt_family="ERVK"
repeat_region 61705..61913

```

Query Match 92.0%; Score 18.4; DB 10; Length 219368;

Best Local Similarity 95.0%; Pred.No. 73;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACGAGCCTCTCCTTTCT 20

|||||

Db 89417 GACGAGCCTCTCCTTTCT 89436

RESULT 4

AL512306

LOCUS

DEFINITION

Human DNA sequence from clone RP11-430C7 on chromosome 1, complete

sequence.

AL512306

VERSION

AL512306.16 GI:18491332

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 133984)

AUTHORS

Bray-Allen,S.

TITLE

Direct Submission

JOURNAL

Submitted (01-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Feb 4, 2002 this sequence version replaced gi:18477311.

COMMENT

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; information on the WormPep database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>  
RP11-430C7 is from the library RPC1-11.2 constructed by the group of Pister de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-430C7 it may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-430C7 is at 133984 in this sequence. The true left end of clone RP11-739N20 is at 97715 in this sequence. The true right end of clone RP11-2317 is at 2003 in this sequence.

#### FEATURES

Location/Qualifiers  
1..133984  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-430C7"  
/clone\_lib="RPC1-11.2"  
123904..123933

#### misc\_feature

/note="Single clone region. Reads generated from a transposon library derived from a single pUC clone. Restriction digest data confirm the assembly."

#### ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 133984;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CGAGCGCTCTCCTTCTCT 20

Db 31667 CGAGCGCTCTCCTTCTCT 31684

#### RESULT 5

#### AC021462

LOCUS AC021462 153023 bp DNA linear HTG 03-APR-2000  
DEFINITION Homo sapiens clone RP11-276C1, WORKING DRAFT SEQUENCE, 21 unordered pieces.

ACCESSION AC021462

VERSION AC021462.3 GI:7387343

KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 153023)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

AUTHORS

Homo sapiens, clone RP11-276C1

Unpublished

REFERENCE 2 (bases 1 to 153023)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A., Coppel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J., Landers,T., Lechoczky,J., Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,N., McSwan,P., McGurk,A., McKernan,K., McPheters,R., Melldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 3, 2000 this sequence version replaced gi:7230200.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L5201

Center clone name: 276 C1

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 141815 bases at least Q40

Consensus quality: 147368 bases at least Q30

Consensus quality: 149481 bases at least Q20

Insert size: 157000; agarose-fp

Insert size: 151023; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1015: contig of 1015 bp in length  
\* 1016 1115: gap of 100 bp  
\* 1116 2276: contig of 1161 bp in length  
\* 2277 2376: gap of 100 bp  
\* 2377 4003: contig of 1633 bp in length  
\* 4010 4110: gap of 100 bp  
\* 4110 5737: contig of 1628 bp in length  
\* 5738 5837: gap of 100 bp  
\* 5838 6939: contig of 1102 bp in length  
\* 6940 7039: gap of 100 bp  
\* 7040 9504: contig of 2465 bp in length  
\* 9505 9604: gap of 100 bp  
\* 9605 11734: contig of 2130 bp in length  
\* 11735 11834: gap of 100 bp  
\* 11835 14111: contig of 2277 bp in length  
\* 14112 14211: gap of 100 bp  
\* 14212 16973: contig of 2762 bp in length  
\* 16974 17073: gap of 100 bp  
\* 17074 19588: contig of 2515 bp in length  
\* 19589 19688: gap of 100 bp  
\* 19689 22275: contig of 2587 bp in length  
\* 22276 22375: gap of 100 bp

```

* 22376 25823: contig of 3448 bp in length
* 25824 25923: gap of 100 bp
* 25924 31307: contig of 5364 bp in length
* 31308 31407: gap of 100 bp
* 31408 34863: contig of 3456 bp in length
* 34864 34963: gap of 100 bp
* 34964 41382: contig of 6419 bp in length
* 41383 41482: gap of 100 bp
* 41483 49024: contig of 7542 bp in length
* 49025 49124: gap of 100 bp
* 49125 58672: contig of 9548 bp in length
* 58673 58772: gap of 100 bp
* 58773 69622: contig of 10850 bp in length
* 69623 69722: gap of 100 bp
* 69723 88191: contig of 18469 bp in length
* 88192 88291: gap of 100 bp
* 88292 107084: contig of 18793 bp in length
* 107085 107184: gap of 100 bp
* 107185 153023: contig of 45839 bp in length.

```

## FEATURES

source

```

1..153023
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-276C1"
/clone_lib="RPC1-11 Human Male BAC"

```

```

1..1015
/note="assembly_fragment"
1116..2276
/note="assembly_fragment"
2377..4009
/note="assembly_fragment"
4110..5737
/note="assembly_fragment"
5838..6939
/note="assembly_fragment"
7040..9504
/note="assembly_fragment"
9605..11734
/note="assembly_fragment"
11835..14111
/note="assembly_fragment"
14212..16373
/note="assembly_fragment"
17074..19588
/note="assembly_fragment"
19689..22275
/note="assembly_fragment"
22376..25823
/note="assembly_fragment"
25924..31307
/note="assembly_fragment"
31408..34863
/note="assembly_fragment"
34964..41382
/note="assembly_fragment"
41483..49024
/note="assembly_fragment"
49125..58672
/note="assembly_fragment"
58773..69622
/note="assembly_fragment"
69723..88191
/note="assembly_fragment"
88292..107084
/note="assembly_fragment"
107185..153023
/note="assembly_fragment"
clone_end:T7
vector_side:left

```

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 153023;

```

Best Local Similarity 100.0%; Pred.No.1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CGACGCCCTCTCCTTCCT 20
   |||||
Db 135254 CGACGCCCTCTCCTTCCT 135271

```

## RESULT 6

CER186/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

/standard_name="R186.6"
/notes="contains similarity to Pfam domain: PF00581
(Rhodanese-like domain), Score=42.7, E-value=2.7e-09, N=1"
/codon_start=1
/product="Hypothetical protein R186.6"
/protein_id="CAB76723.1"
/db_xref="GI:7160720"
/translation="MESSLSKSDMMENTNHLITLITKRKICILEAASGDEAKSRD
AFSDVHIESRLFLFHSNLSHAGVEHFLQFQFARSGQINDCHVIVIDRGQMWSSY
ITVFKLFGKQVLSLGGVIGKMTQHQARSGQYKTEGDAPRXPQGFELASWDSVI
ITVDVLLNEIDNFVDQAQTKDEFLGTAGALYGHKARNIPDVAVDWAGQWK
DADHLKGLFNKALSLEKPVVYCVIVISQNFDPKS"
complement(join(2877..3009,3322..3577,3621..3700,
4050..4281,4342..4471))
/gene="R186.1"
complement(join(2877..3009,3322..3577,3621..3700,
4050..4281,4342..4471))
/gene="R186.1"
/standard_name="R186.1"
/notes="Weak similarity to Human T10 mRNA product
(TR:G40G263)
CDNA EST Yk235e12.5 comes from this gene
CDNA EST Yk734f3.5 comes from this gene
CDNA EST Yk734f3.3 comes from this gene
CDNA EST Yk795b01.5 comes from this gene
CDNA EST Yk235e12.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein R186.1"
/protein_id="CAB01439.1"
/db_xref="GI:3879196"
/translation="MCITFTKAKSLDKYKLLNNRDLRLPTKEMWHGIIISG
VDQENARGFWLGLNENGRIGMLSITQTOESKNLHAPSGGIVNSELNANDSKME
SLKCKAKYNGFOLVAVERKSTGLYEVRTLQANQVDEIVCQKDEYHVNSNPPTKP
YQKAVQCKKLRLHLENSDQFSDVDFEKLKLSAKNTTQWYDQAQYQYQYQYVNEYNR
FLSAIFTKYPEGRVTGRTCTHTLTVDQKINILERRLLPEQSTWHDASFVFLNGS"
join(5527..5856,6303..6443,6489..6682,6764..6921,
6965..7152)
/gene="srp-35"
join(5527..5856,6303..6443,6489..6682,6764..6921,
6965..7152)
/gene="srp-35"
/standard_name="R186.2"
/codon_start=1
/product="C. elegans SRD-35 protein (corresponding
sequence R186.2)"
/protein_id="CAB01441.1"
/db_xref="GI:3879198"
/translation="MVPQLNVDSDPSKIVEINEIYPIFLITLIFHSVLVFIIF
FSPSHLNF.KFLILITSLDVFATSLFVIOPRILSEYOVNVPYCYGCPRENTQFC
PSLFMMWQTVSIAVGISLTLYLKYKHLKASKPLGKWLVRSLFLYIFCLVSMCCA
ZVVVLNVLPLNEBKESFSDTLEFKYGNMKLGNLPLNFAMIAIAYLCPSPITTF
WFKWTKNSLNNALSGSPYLRHAKNVMGITIQQVMHFIYFIPFTLYSYSLTGT
KIFVQFFMAWSPNLAAFDPLNLVFPVYRTKIKSWFGKQQTSTIRIASLTFSG
YF"
complement(join(7346..7427,7474..7673,7776..7886,
7938..8071,8121..8223,8267..8359))
/gene="R186.3"
complement(join(7346..7427,7474..7673,7776..7886,
7938..8071,8121..8223,8267..8359))
/gene="R186.3"
/standard_name="R186.3"
/notes="Similarity to Mouse signal recognition particle
receptor beta subunit (SW:SRPB_MOUSE)
CDNA EST Yk522a4.3 comes from this gene
CDNA EST Yk522a4.5 comes from this gene
CDNA EST Yk807b04.3 comes from this gene
CDNA EST Yk807b04.5 comes from this gene
CDNA EST Yk1275a10.5 comes from this gene
CDNA EST Yk1275a10.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein R186.3"
/protein_id="CAB01443.1"

```

```

/db_xref="GI:3879200"
/translation="MDKIDFNDFTTAVLATAVIGLTVLLLVLKSPASNKRVLPV
GLMKCKTFTFOLSQKEAEPYTKTKYTSMNENKITLRKKEKEIIDIYDGNDRURQ
KIENHLRSGLRIVFVVDSAFNRDVAELFYVALENKDYKVPILIACHKQDLS
LAKTEKVIKNSLEKEIGLKNKRAALIGTGSEKRSILTDTGIDFKWEDLKQEVs
FVSTKSNBDPFGVHEIASFVRA"
complement(join(8532..8626,8775..8878,8928..8975,
9541..9644))
/gene="R186.8"
complement(join(8532..8626,8775..8878,8928..8975,
9541..9644))
/gene="R186.8"
/standard_name="R186.8"
/codon_start=1
/product="Hypothetical protein R186.8"
/protein_id="CAE17911.1"
/db_xref="GI:33300311"
/translation="MARGIHLQHVPTKELVSTNNGLKCKILKELYSFKLPLLQYK
VMGAKSYVIVOLASVITGSTRVWVZERAERFAGIFYDPAYGKSCLFEEVKKVKS
EUPKIRGIYSIEN"
join(10042..10065,10127..10307,10591..10807,10916..11285,
11345..11728)
/gene="lin-46"
join(10042..10065,10127..10307,10591..10807,10916..11285,
11345..11728)
/gene="lin-46"
/standard_name="R186.4"
/notes="Similarity to Rat gephyrin (SW:GEPH_RAT), contains
similarity to Pfam domain: PF00994 (Molybdenum cofactor
biosynthesis protein), Scores=82.9, E-value=4.4e-10, N=1
CDNA EST CERKS51FB comes from this gene
CDNA EST AF042069 comes from this gene
CDNA EST Yk1087f07.5 comes from this gene
CDNA EST Yk1087f07.3 comes from this gene"
/codon_start=1
/product="C. elegans LIN-46 protein (corresponding
sequence R186.4)"
/protein_id="CAB01440.2"
/db_xref="GI:14530531"
/translation="MSSSGLLKPATLDDVQKLEDLCKLFPQPKETVNTVSLTKGRIL
AEDITTEYDPAQTSIVDGFPAIIVNQLGTXREIVGLSTAVTPYNAELISNECVRTI
GGVDPGADTVVPIENVALKEKCEIVLRKPKGNIREVSGSEAKTGLDKDGHHL
DWSITLLEALGISOVEIYKPRVCVLVTSIGLSNKNMYGSPNRSOLLBLFQSGGFTA
IDAGSSTEHTIETEVEEKIRTPAACPACVLTVGGCAQVIREVAKTLKFKFEIQQDVPSTGN
FVYSTKIDETPVLVLSIPFEPYVSSWIGANLFPSPILRAMGONSETSRFKALLQFP
IKSTSETFLRARSEVSKGNLSTPLGCEDFPANSILEVKNTCSFAGDVVDLRFPA"
join(12268..12365,12414..12566,12619..12982,13027..13258,
13928..14160)
/gene="R186.7"
join(12268..12365,12414..12566,12619..12982,13027..13258,
13928..14160)
/gene="R186.7"
/standard_name="R186.7"
/notes="CDNA EST Yk417a1.3 comes from this gene
CDNA EST Yk424g3.3 comes from this gene
CDNA EST Yk497g9.3 comes from this gene
CDNA EST Yk290a11.5 comes from this gene
CDNA EST Yk417a1.5 comes from this gene
CDNA EST Yk424g3.5 comes from this gene
CDNA EST Yk499a4.5 comes from this gene
CDNA EST Yk499a4.3 comes from this gene
CDNA EST Yk517d12.5 comes from this gene
CDNA EST Yk517d12.3 comes from this gene
CDNA EST Yk522f6.5 comes from this gene"
Query Match 87.0%; Score 17.4; DB 3; Length 29134;
Best Local Similarity 94.7%; Pred. No. 2,5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ACACGCCCTCTCTCTTTCCT 20
|||||
Db 18201 ACACGCCCTCTCTCTTTCGT 18183

```



```

REFERENCE
AUTHORS Siebers B.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1997) B. Siebers, FB 9 Mikrobiologie,
Universitaet-GH Essen, Universitaetsstrasse 5, 45117 Essen, FRG
FEATURES
source
1..1014
/organism="Thermoproteus tenax"
/mol_type="genomic DNA"
/strain="Kral (DSM 2078)"
/db_xref="taxon:2271"
1..1014
/gene="pfp"
1..1014
/gene="pfp"
/EC_number="2.7.1.90"
/feature="pyrophosphate--fructose-6-phosphate
1-phosphotransferase"
/codon_start=1
/transl_table=1
/product="alternative product name:
pyrophosphate-dependent phosphofructokinase"
/protein_id="CAA74985.1"
/db_xref="GI:3087896"
/db_xref="GOA:O57694"
/db_xref="SPTREML:O57694"
/translation="MKIGLVLTGGDAGFLNATVTFVKLARKEHVAIVHGRGLLN
KEVRSRLDLPDFSGGTVIRTSNPFKDERARLLSNKELGLDVVAIGDD
TLGAAGRAQRGILDVGIPTKIDNDVYGTDTIGFDSAVNAAEATESFTLLISHE
RIGVYVMGREGAIWALTGLSTWADAVLIPERPAWDSVAKRVEAYNERMALVIV
SEGLIKYGGPKDEYGHRLGGVGNELABYERSTGEARAVLGHITRGVPTAFDRI
LAVRYATAAYEAVENGRYGVWVAYNSNGDIAPVPIVDVVGKRLVSGYWMRLYETWFD
LAG"

ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 1014;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCTCTCCTTCCT 20
|||||
Db 495 GACGACGCTATCCTTCGT 476

RESULT 9
CHKPLB
LOCUS Chicken cardiac phospholamban (plb) mRNA linear VRT 06-MAR-1995
DEFINITION
ACCESSION M59039
VERSION M59039.1 GI:212575
KEYWORDS phospholamban.
SOURCE Gallus gallus (chicken)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 3312)
Toyofuku,T. and Zak,R.
AUTHORS
TITLE Characterization of cDNA and genomic sequences encoding a chicken
phospholamban
JOURNAL J. Biol. Chem. 266 (9), 5375-5383 (1991)
MEDLINE 91170195
PUBMED 1825996
COMMENT Original source text: Chicken (broiler breeders) 7-week old adult
heart, cDNA to mRNA, clones CPL-[6,12,15].
FEATURES
source
1..3312
Location/Qualifiers
1..3312
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="broiler breeder"
/db_xref="taxon:9031"
/clone="CPL-15"
/clone="CPL-6"

Query Match 84.0%; Score 16.8; DB 5; Length 3312;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCTCTCCTTCCT 20
|||||
Db 183 GACAACTCCTCTCCTTCCT 202

RESULT 10
CHKPLB2
LOCUS Chicken phospholamban (plb) gene, exon 2. linear VRT 07-MAR-1995
DEFINITION
ACCESSION M59038
VERSION M59038.1 GI:212578
KEYWORDS phospholamban.
SEGMENT 2 of 2
SOURCE Gallus gallus (chicken)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 3423)
Toyofuku,T. and Zak,R.
AUTHORS
TITLE Characterization of cDNA and genomic sequences encoding a chicken
phospholamban
JOURNAL J. Biol. Chem. 266 (9), 5375-5383 (1991)
MEDLINE 91170195
PUBMED 1825996
COMMENT Original source text: Chicken (White leghorn) adult liver DNA,
clones GPL-1, and GPL-2.
FEATURES
source
1..3423
Location/Qualifiers
1..3423
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="White leghorn"
/sub_species="domesticus"
/db_xref="taxon:9031"
/clone="GPL-1"
/tissue_type="liver"
/dev_stage="adult"

```

```

gene      /tissue_lib="pEMBL3 chicken genomic"
join(M59037.1:1511..1803,1..3423)
mRNA      /gene="plb"
join(M59037.1:1621..1745,89..3423)
mRNA      /gene="plb"
product="phospholamban"
join(M59037.1:1621..1745,89..3276)
mRNA      /gene="plb"
product="phospholamban"
join(M59037.1:1745,89..951)
mRNA      /gene="plb"
product="phospholamban"
join(M59037.1:1621..1745,89..567)
intron    /product="phospholamban"
order(M59037.1:1746..1803,1..88)
exon      /gene="plb"
89..3423
exon      /gene="plb"
number=2
89..3276
exon      /gene="plb"
number=2
89..951
exon      /gene="plb"
number=2
89..567
CDS       /gene="plb"
number=2
209..367
polyA_signal /codon_start=1
product="phospholamban"
protein_id="AA63167.1"
db_xref="GI:212580"
translation="MEKQYITRSALRASTLEVNVPQRLQELFVNFCILICLLI
ICIVMLL"
polyA_signal 534..539
polyA_signal /gene="plb"
930..935
polyA_signal /gene="plb"
3254..3259
polyA_signal /gene="plb"

```

## ORIGIN

```

Query Match      84.0%; Score 15.8; DB 5; Length 3423;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GACGACGCTCTCCTTCTCT 20
Db 146 GACAACTCCTCTCCTTCTCT 165
|||||
146 GACAACTCCTCTCCTTCTCT

```

```

RESULT 11
AK065803/c
LOCUS      AK065803      3491 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oriza sativa (japonica cultivar-group) cDNA clone:J013036D01, full
            insert sequence.
VERSION    AK065803.1 GI:32975821
KEYWORDS   FLI_CDNA; CAP trapper.
SOURCE     Oriza sativa (japonica cultivar-group)
ORGANISM   Oriza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzoae; Oryza.
REFERENCE  1
            The Rice Full-length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-length cDNA Project team:
            Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
            Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I.,
            Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

```

Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 3491)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kaga, I., Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuki, Y., Ryu, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shishiki, T., Sogabe, Y., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

## Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp).

Tel: 81-29-838-7007, Fax: 81-29-838-7007

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Itoh, M., Kaga, I., Kawai, J., Imamura, K., Imotani, K., Ishii, Y., Kishikawa-Hirozane, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

## TITLE

## JOURNAL

## COMMENT





```

88 14.8 74.0 512 13 US-10-027-632-194334 Sequence 194334,
89 14.8 74.0 512 13 US-10-027-632-194335 Sequence 194335,
90 14.8 74.0 512 16 US-10-027-632-194334 Sequence 194334,
91 14.8 74.0 512 16 US-10-027-632-194335 Sequence 194335,
92 14.8 74.0 524 9 US-09-974-300-7215 Sequence 7215, Ap
93 14.8 74.0 525 13 US-10-424-599-41533 Sequence 41533, A
94 14.8 74.0 531 15 US-10-029-386-6226 Sequence 6226, Ap
95 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
96 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
97 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
98 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
99 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
100 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
101 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
102 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
103 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
104 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
105 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
106 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
107 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
108 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
109 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
110 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
111 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
112 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
113 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
114 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
115 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
116 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
117 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
118 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
119 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
120 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
121 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
122 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
123 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
124 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
125 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
126 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
127 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
128 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
129 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
130 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
131 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
132 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
133 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
134 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
135 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
136 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
137 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
138 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
139 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
140 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
141 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
142 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
143 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
144 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
145 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
146 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
147 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
148 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
149 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
150 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
151 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
152 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
153 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
154 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
155 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
156 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
157 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
158 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
159 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,
160 14.8 74.0 531 15 US-10-027-632-320214 Sequence 320214,

```

## ALIGNMENTS

## RESULT 1

```

US-09-877-819B-5/c
; Sequence 5, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, Scott
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address tag
US-09-877-819B-5

```

Query Match 100.0%; Score 20; DB 10; Length 20;

Best Local similarity 100.0%; Pred. No. 5.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGAGCGCTCTCTCTCTCTCT 20

Db 20 GAGGAGCGCTCTCTCTCTCTCT 1

## RESULT 2

```

US-09-877-819B-6
; Sequence 6, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, Scott
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

```

```
; OTHER INFORMATION: Capture tag
US-09-877-819B-6

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGCGCTCTCCTTTCTCT 20
   |||||
Db 1 GAGGAGCGCTCTCCTTTCTCT 20

RESULT 3
US-09-877-819B-33
; Sequence 33, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: White, Scott
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 33
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Capture sequence
US-09-877-819B-33

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGCGCTCTCCTTTCTCT 20
   |||||
Db 1 GAGGAGCGCTCTCCTTTCTCT 20

RESULT 4
US-09-877-819B-35/c
; Sequence 35, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: White, Scott
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 35
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address sequence
US-09-877-819B-35

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGCGCTCTCCTTTCTCT 20
   |||||
Db 20 GAGGAGCGCTCTCCTTTCTCT 1

RESULT 5
US-09-877-819B-35

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGCGCTCTCCTTTCTCT 20
   |||||
Db 20 GAGGAGCGCTCTCCTTTCTCT 1

RESULT 6
US-10-425-114-15395/c
; Sequence 15395, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jirongdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15395
; LENGTH: 880
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LI33059-048-G11_FLI
US-10-425-114-15395

Query Match      82.0%; Score 16.4; DB 13; Length 880;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CGACGCGCTCTCCTTTCTCT 20
   |||||
Db 64 CGCGCGCTCTCCTTTCTCT 47

RESULT 7
US-10-087-192-1891
; Sequence 1891, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
```

```
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1891
; LENGTH: 37923
; TYPE: DNA
; ORGANISM: Mus musculus
; NAME/KEY: misc feature
; LOCATION: (1)...(37923)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1891

Query Match      82.0%; Score 16.4; DB 13; Length 37923;
Best Local Similarity 94.4%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACGAGCGCTCTCCTTTC 18
Db 24589 GACGAGCGCTCTCCTTGC 24506

RESULT 8
US-09-910-082A-309/c
; Sequence 309, Application US/99910082A
; Publication No. US20030119731A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Balcomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Watkins, Maren
; APPLICANT: Garrett, James E.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; TITLE OF INVENTION: Omega-Concept,des
; FILE REFERENCE: 2314-241
; CURRENT APPLICATION NUMBER: US/09/910,082A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 60/219,616
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/265,888
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 309
; LENGTH: 307
; TYPE: DNA
; ORGANISM: Conus rattus
US-09-910-082A-309

Query Match      79.0%; Score 15.8; DB 10; Length 307;
Best Local Similarity 89.5%; Pred. No. 4.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ACGAGCGCTCTCCTTTCCT 20
Db 105 ACGAGCGCTCTCCTTACT 87

RESULT 9
US-10-079-623-208
; Sequence 208, Application US/10079623
; Publication No. US20020169302A1
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; TITLE OF INVENTION: Compositions isolated from bovine
; FILE REFERENCE: 11000.1044G3
; CURRENT APPLICATION NUMBER: US/10/079,623
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Bovine
; OTHER INFORMATION: n = A,T,C or G
US-10-079-623-208

Query Match      79.0%; Score 15.8; DB 14; Length 402;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGAGCGCTCTCCTTTC 19
Db 375 GAGGAGCGCTCTCCTTCC 393

RESULT 10
US-10-425-114-27257
; Sequence 27257, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27257
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4729-021-F1_FLI
US-10-425-114-27257

Query Match      79.0%; Score 15.8; DB 13; Length 1007;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGAGCGCTCTCCTTTC 19
Db 142 GCGCGCGCTCTCCTTCC 160

RESULT 11
US-10-424-599-99564
; Sequence 99564, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
```

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 99564  
; LENGTH: 1207  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_60921C.1  
US-10-424-599-99564

Query Match 79.0%; Score 15.8; DB 13; Length 1207;  
Best Local Similarity 89.5%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACAGCGCTCTCTCTTCT 20  
||||| ||||| ||||| |||||  
Db 27 ACAGCGACTTCTCTTCT 45

RESULT 12  
US-09-974-300-234/c  
; Sequence 234, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berk, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: 10085,500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 234  
; LENGTH: 1362  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-234

Query Match 79.0%; Score 15.8; DB 9; Length 1362;  
Best Local Similarity 89.5%; Pred. No. 3.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGCGCTCTCTCTTCT 19  
||||| ||||| ||||| |||||  
Db 102 GAGCGCTCTCTCTTCT 84

RESULT 13  
US-10-425-114-1449/c  
; Sequence 1449, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 1449  
; LENGTH: 1588  
; TYPE: DNA  
; ORGANISM: Zea mays

; FEATURE:  
; OTHER INFORMATION: Clone ID: 700153483\_FLI  
US-10-425-114-1449

Query Match 79.0%; Score 15.8; DB 13; Length 1588;  
Best Local Similarity 89.5%; Pred. No. 3.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGCGCTCTCTCTTCT 19  
||||| ||||| ||||| |||||  
Db 153 GAGCGACTCGCTTCT 135

RESULT 14  
US-10-311-455-1095/c  
; Sequence 1095, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determini  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: JS/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1095  
; LENGTH: 7690  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1095

Query Match 79.0%; Score 15.8; DB 15; Length 7690;  
Best Local Similarity 89.5%; Pred. No. 3.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACAGCGCTCTCTCTTCT 20  
||||| ||||| ||||| |||||  
Db 7549 ACAGTCTCTCTCTTCT 7531

RESULT 15  
US-10-311-455-198/c  
; Sequence 198, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determini  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 198  
; LENGTH: 14924  
; TYPE: DNA

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 14698, 14712, 14714...14715, 14717
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-198

```

```

Query Match          79.0%; Score 15.8; DB 15; Length 14924;
Best Local Similarity 89.5%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 2 ACAGCGCTCTCCTTTCT 20
Db 12332 ACAACACCTCTCCTTTCT 12314

```

```

RESULT 16
US-10-240-452-22/c
; Sequence 22, Application US/10240452
; Publication No. US20030162194A1
; GENERAL INFORMATION:
; APPLICANT: OLEX, Alexander
; APPLICANT: PIERPNEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Apoptosis
; FILE REFERENCE: 5013.1006
; CURRENT APPLICATION NUMBER: US/10/240,452
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03969
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 78
; SEQ ID NO 22
; LENGTH: 14924
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (14698, 14712, 14714...14715, 14717)
US-10-240-452-22

```

```

Query Match          79.0%; Score 15.8; DB 15; Length 14924;
Best Local Similarity 89.5%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 2 ACAGCGCTCTCCTTTCT 20
Db 12332 ACAACACCTCTCCTTTCT 12314

```

```

RESULT 17
US-10-087-192-1057
; Sequence 1057, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01

```

```

; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1057
; LENGTH: 71953
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(71953)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1057

```

```

Query Match          79.0%; Score 15.8; DB 13; Length 71953;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 2 ACAGCGCTCTCCTTTCT 20
Db 46115 ATGACACCTCTCCTTTCT 46133

```

```

RESULT 18
US-10-087-192-226
; Sequence 226, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226
; LENGTH: 653122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(653122)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-226

```

```

Query Match          79.0%; Score 15.8; DB 13; Length 653122;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 2 ACAGCGCTCTCCTTTCT 20
Db 251537 ACCAGCCTCTCCTTTCT 251555

```

```

RESULT 19
US-10-085-783A-3539
; Sequence 3539, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1272.82 Seconds

(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819b-33

Perfect score: 20

Sequence: 1 gacgagcctctctcttct 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 180 summaries

Database :

RST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rpd:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gssl:\*

29: gb\_gssl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	18.4	92.0	733	29	BZ967818
2	17.4	87.0	379	9	AB076957
3	17.4	87.0	503	14	CD111750
4	17.4	87.0	695	28	BH705590

BH718652	BOMM90TR
BB813656	BB813656
BF622509	FVSMEO00
BX539720	Leishmani
BZ894876	Hg4_0166
CC483605	CH240_312
BU214159	603109140
BX276312	BX276312
BU107606	603111219
BZ349842	hr46d12.5
BX276311	BX276311
BU108783	603109972
BF224860	uz11g12.x
BU289346	604165581
AL163994	Tetraodon
CD430028	ETH1_13.F
BM489672	pgm2n.pk0
BX539026	Leishmani
CG331413	CGXDG51TV
BU143208	603126266
BU144044	603228840
CG331399	CGXDG51TH
BI521933	603081256
AG075431	Pan trogl
BI831377	603074366
BU144412	603229615
CC393953	FURKGA5TB
BU346064	604172642
BU109258	603112171
BU144375	603229576
BU145093	603229423
CA764236	AP53-Rpf
BU144698	603229629
BU144278	603229808
CG289248	CGXDI83TV
BU143249	603229230
CD241445	AGENCOURT
BU106836	603110686
BU144429	603230184
BU369135	603598625
BG343181	HVSMEO00
BU144543	603228913
BU143936	603230222
BU108727	603118293
AL098359	Drosophil
BU506565	AGENCOURT
AL188400	Tetraodon
AL347118	Tetraodon
BC032063	Homo sapi
BB184379	BB184379
BS569276	BQGV37TF
CD296086	Strpue91.
BS538515	Leishmani
CF748983	UI-M-HUO-
CC605307	QGVPP16TV
BZ216439	CH230-323
BH585750	BQGPL84TF
BG660059	102406731
BX464371	BX464371
BI336789	AR081D12B
BB071518	BB071518
BM744826	K-EST0018
BM737808	K-EST0001
CF938309	NcESTgab3
CD667452	NcEST3b92
CF943655	NcEST3b92
AZ004718	RPCI-23-3
BZ346996	602021969
BF688771	602184936
BQ535248	LEAF3_9.F
CE439441	tigr-gss-
BE600743	PII_89.F1
BF651869	274762.MA

[illegible]



## ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 733;  
 Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGAGCCTCTCTCTTCTCT 20  
 |||||  
 Db 706 GACGAGCTCTCTCTCTTCTCT 725

## RESULT 2

AB076957 379 bp mRNA linear EST 02-JUL-2002  
 DEFINITION AB076957 Human vestibular cDNA library Homo sapiens cDNA clone  
 402V5-12-54, mRNA sequence.  
 ACCESSION AB076957  
 VERSION AB076957.1 GI:21678535  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 379)  
 AUTHORS Abe, S., Koyama, K., Usami, S. and Nakamura, Y.  
 TITLE Construction of a vestibular-specific cDNA library  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Satoko Abe  
 Institute of Medical Science  
 The University of Tokyo, Human Genome Center  
 4-6-1, Minato-Ku, Tokyo 108-8639, Japan  
 Tel: 81-3-5449-5375  
 Fax: 81-3-5449-5406  
 Email: satoko@ims.u-tokyo.ac.jp.  
 Location/Qualifiers  
 1..379  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="402V5-12-54"  
 /tissue\_type="vestibule"  
 /clone\_lib="Human vestibular cDNA library"

## FEATURES

source  
 1..379  
 422 GACGAGCCTCTCTCTTCTCT 440

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 379;  
 Best Local Similarity 94.7%; Pred. No. 3.2e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAGCCTCTCTCTTCTCT 20  
 |||||  
 Db 19 ACGAGCCTCTCTCTTCTCT 37

## RESULT 3

CD111750 503 bp mRNA linear EST 14-SEP-2003  
 LOCUS ME1-0021T-D051-E02-U.B ME1-0021 Schistosoma mansoni cDNA clone  
 DEFINITION ME1-0021T-D051-E02-E, mRNA sequence.  
 ACCESSION CD111750  
 VERSION ME1-0021T-D051-E02-E, mRNA sequence.  
 KEYWORDS EST.  
 SOURCE Schistosoma mansoni  
 ORGANISM Schistosoma mansoni

## REFERENCE

AUTHORS Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.  
 1 (bases 1 to 503)  
 Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,  
 Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,  
 Kicajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,  
 Coulson, P.S., Dillon, G.P., Farias, L.F., Gregorio, S.P., Ho, P.L.,  
 Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,  
 Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,  
 Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,

Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,  
 Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.  
 Transcriptional analysis of the acelomate human parasite Schistosoma  
 mansoni  
 Nat. Genet. 35 (2), 148-157 (2003)  
 22879926  
 Contact: Dr. Sergio Verjovski-Almeida  
 Departamento de Bioquímica  
 Instituto de Química - Universidade de São Paulo  
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
 Brasil  
 Tel: +55-11-3091-2173  
 Fax: +55-11-3091-2186  
 Email: verj@iq.usp.br  
 This sequence was derived from the FAPESP Schistosoma mansoni EST  
 Genome Project. All sequences in the project were assembled and  
 annotated. This entry and all the assembled sequences can be seen  
 in the following URL <http://bioinfo.iq.usp.br/schisto/>  
 Plate: ME1-0021T-D051 row: 2 column: E.

## FEATURES

source  
 1..503  
 /organism="Schistosoma mansoni"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6183"  
 /clone="ME1-0021T-D051-E02-E"  
 /sex="mixed pool"  
 /dev\_stage="egg"  
 /lab\_host="MUS musculus"  
 /clone\_lib="ME1-0021"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 503;  
 Best Local Similarity 94.7%; Pred. No. 3.5e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGAGCCTCTCTCTTCTCT 19  
 |||||  
 Db 422 GACGAGCCTCTCTCTTCTCT 440

## RESULT 4

BH705590 695 bp DNA linear GSS 20-FEB-2002  
 LOCUS BOMME07TR BO.2.3 KB Brassica oleracea genomic clone BOMME07,  
 DEFINITION BOMME07TR BO.2.3 KB Brassica oleracea genomic clone BOMME07,  
 genomic survey sequence.  
 ACCSSION BH705590  
 VERSION BH705590.1 GI:18788415  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 695)  
 AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.  
 TITLE Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other\_GSSs: BOMME07TF  
 Contact: Chris Town

## FEATURES

source  
 1..695  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TC1000DH3"  
 /db\_xref="taxon:3712"

/clone="BOMMEC7"  
 /clone\_lib="BO 2 3 X5"  
 /note="vector: pHOSt; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 695;  
 Best Local Similarity 94.7%; Pred. No. 3.7e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGCAGCGCTCTCCTTTCC 19  
 |||||  
 Db 5 GAGCAGCGCTCTCCTTTCC 23

## RESULT 5

BH718652  
 LOCUS BH718652 773 bp DNA linear GSS 20-FEB-2002  
 DEFINITION BOMMY90TR BO 2 3 KB Brassica oleracea genomic clone BOMMY90,  
 genomic survey sequence.

ACCESSION BH718652  
 VERSION BH718652.1 GI:16817384  
 KEYWORDS GSS.

## SOURCE

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica;

1 (bases 1 to 773)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other GSSs: BOMMY90TF

Contact: Chris Town

TIGR Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: Sheared ends.

## FEATURES

source

1..773  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOMMY90"  
 /clone\_lib="BO 2 3 KB"  
 /note="vector: pHOSt; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 773;  
 Best Local Similarity 94.7%; Pred. No. 3.8e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGCAGCGCTCTCCTTTCC 19  
 |||||  
 Db 464 GAGCAGCGCTCTCCTTTCC 482

## RESULT 6

BH813656/c  
 LOCUS BH813656 358 bp mRNA linear EST 19-NOV-2001  
 DEFINITION BH813656 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus  
 musculus cDNA clone G730020C03 3', mRNA sequence.

ACCESSION BH813656  
 VERSION BH813656.1 GI:16986285  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 358)  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, C., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shigesawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akai, S., Tanaka, T., Tomaru, A., Toyota, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

## JOURNAL

COMMENT

Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

## FEATURES

source

1..358  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G730020C03"  
 /tissue\_type="lung"  
 /cell\_line="RCB-0558 LLC"  
 /clone\_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 358;  
 Best Local Similarity 90.0%; Pred. No. 5.8e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGCAGCGCTCTCCTTTCT 20  
 |||||  
 Db 82 GTCCAGCGCTCTCCTTTCT 53

## RESULT 7

BF622509  
 LOCUS BF622509 475 bp mRNA linear EST 17-OCT-2001  
 DEFINITION HVSMea0004K15f Hordeum vulgare seedling shoot EST library  
 HVCDNA0001 (Cold stress) Hordeum vulgare subsp. vulgare cDNA clone  
 HVSMea0004K15f, mRNA sequence.

## ACCESSION

BF622509

## VERSION

BF622509.2 GI:113081086

## ORIGIN

/clone.lib="Hg pUC18 Library"  
/note=Vector: pUC18; Site 1: SmaI; A shotgun library was  
constructed from Halobaculum gomorrense genomic DNA using  
pUC18/SmaI/BAP plasmid"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 28; Length 530;  
Best Local Similarity 90.0%; Pred. No. 6.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCCTCTCCTTCCT 20  
||||| ||||| ||||| |||||  
Db 247 GACGACGCCTCTCCTTCCT 228

## RESULT 10

CC483605

LOCUS

DEFINITION CH240\_312D19.77 CHORI-240 Bos taurus genomic clone CH240\_312D19,  
genomic survey sequence.

ACCESSION

CC483605

VERSION

CC483605.1

KEYWORDS

GSS.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,  
Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M.,  
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,  
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,  
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,  
Dalrymple, B.P. and Tellam, R.

TITLE

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL

Unpublished (2003)

COMMENT

Other\_GSSs: CH240\_312D19.FARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4R6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rnoit@ccgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering\_information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 312 row: D column: 19

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..551

/organism="Bos taurus"

/mol\_type="genomic DNA"

/strain="breed: Hereford"

/db\_xref="taxon:9913"

/clone="CH240\_312D19"

/sex="Male"

/cell\_type="Blood"

/clone.lib="CHORI-240"

/note=Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull LI Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 551;

Best Local Similarity 90.0%; Pred. No. 6.4e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCCTCTCCTTCCT 20  
||||| ||||| ||||| |||||  
Db 529 GACGACGCCTCTCCTTCCT 548

## RESULT 11

BU214159

LOCUS

DEFINITION

603109140F1 CSEQCHN04

Gallus gallus

sequence.

ACCESSION

BU214159

VERSION

BU214159.1

KEYWORDS

EST.

SOURCE

Gallus gallus (chicken)

ORGANISM

Gallus gallus

REFERENCE

AUTHORS

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
Zong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE

A Comprehensive Collection of Chicken cDNAs

JOURNAL

Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

22335534

PUBMED

12445392

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..554

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="White Leghorn, HiseX"

/db\_xref="taxon:9031"

/clone="CHEST5309"

/tissue\_type="whole embryo"

/dev\_stage="20-21"

/lab\_host="DH10B"

/clone.lib="CSEQCHN04"

/note=Organ: whole embryo; Vector: pBluescript II KS(+);

Site 1: EcoRI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 554;

Best Local Similarity 90.0%; Pred. No. 6.4e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCCTCTCCTTCCT 20  
||||| ||||| ||||| |||||  
Db 145 GACGACGCCTCTCCTTCCT 164

## RESULT 12

BX275312

```

LOCUS      BX276312                570 bp    mRNA    linear    EST 27-FEB-2003
DEFINITION BX276312 AGENAE Gallus gallus multi-tissues normalized library
            (gcag) Gallus gallus cDNA clone gcag008c.e.09 5prim, mRNA
sequence.
ACCESSION  BX276312
VERSION     BX276312.1  GI:28598803
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 570)
AUTHORS   Herault, F., Le Meuth-Metzinger, V., Desert, C., Retout, E., Piumi, F.,
            Klopp, C. and Douaire, M.
TITLE     Construction and primary characterization of chicken normalized
            multi-tissue cDNA libraries
JOURNAL    Unpublished (2003)
COMMENT    Contact: Douaire M
            INRA, UMR INRA-ENSAR Genetique Animale
            65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
            Tel: +33 (0) 2.23.48.54.63
            Fax: +33 (0) 2.23.48.54.70
            Email: Madeleine.Douaire@roazhon.inra.fr
            Sequence cleaned of vector, adaptor and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
            Plate: 0008 row: e column: 9
            Seq primer: M13R.
            Location/Qualifiers
                1..570
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /db_xref="taxon:9031"
                /clone="gcag008c.e.09"
                /tissue_type="multi-tissues"
                /dev_stage="from embryos to adults"
                /lab_host="DH10B"
                /clone_lib="AGENAE"
                /lab="Gcag"
                /note="vector: pYT73D-pac; tissues: brain, embryos,
                kidney, multi-tissues, muscle, pancreas, skin, testis,
                liver, adipose tissue, granulosa, utero-vaginal gland,
                oviduct, small follicle, ovary, hypothalamus, pituitary
                gland, ileon, jejunum, caecum, duodenum, spleen,
                fabricius gland, bone marrow, thymus, hematopoietic
                progenitor cells. Clone distribution : AGENAE Resource
                centre, Francois PIGNI, Francois Piumi.inra.fr, INRA, CEA
                Radiobiologie et Etude du genome (IREG), Domaine de
                Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

FEATURES             source
    source
        1..570
        /organism="Gallus gallus"
        /mol_type="mRNA"
        /strain="White Leghorn, H-sex"
        /db_xref="taxon:9031"
        /clone="CHESF5866"
        /dev_stage="36"
        /lab_host="DH10B"
        /clone_lib="CSEQCHL12"
        /note="Organ: heart; Vector: pBluescript II KS(+); Site 1:
        EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
        [Stratagene] vector to accommodate cDNA produced with the
        T-trimmed protocol (Construction of uni-directionally
        cloned cDNA libraries from messenger RNA for improved 3'
        end DNA sequencing by Glenn Fu, et al. U.S. Patent #
        6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
        Ligase in double stranded adaptor containing BsgI and
        BamHI sites [5'gcccgcgcagcccgatccgcagaaaaag]
        [5'aattcttttttggatccggggtgcagc]"

ORIGIN
    Query Match      84.0%; Score 16.8; DB 13; Length 570;
    Best Local Similarity 90.0%; Pred. No. 6.4e+03;
    Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      1  GACGACGCTCTCCTTCCT 20
        |||||
Db      169 GACAACTCTCTCTCTTCCT 188

RESULT 14
BX249842
LOCUS      hr46d12.g1 WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum
            bicolor genomic clone hr46d12 5', genomic survey sequence.
DEFINITION
ACCESSION  BX249842
VERSION     BZ349842.1  GI:24912187
KEYWORDS   GSS.
SOURCE     Sorghum bicolor (sorghum)
ORGANISM   Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE  1 (bases 1 to 572)
AUTHORS   Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
            Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
            Zutavern, T., Palmer, L., McCombie, W.R. and Martienssen, R.A.
            Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
            Unpublished (2002)
            Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mcombie@cshl.org

```

```

Phasianinae; Gallus.
1 (bases 1 to 572)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickie, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
MEDLINE
PUBMED
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(MIIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
                1..572
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="White Leghorn, H-sex"
                /db_xref="taxon:9031"
                /clone="CHESF5866"
                /dev_stage="36"
                /lab_host="DH10B"
                /clone_lib="CSEQCHL12"
                /note="Organ: heart; Vector: pBluescript II KS(+); Site 1:
                EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
                [Stratagene] vector to accommodate cDNA produced with the
                T-trimmed protocol (Construction of uni-directionally
                cloned cDNA libraries from messenger RNA for improved 3'
                end DNA sequencing by Glenn Fu, et al. U.S. Patent #
                6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
                Ligase in double stranded adaptor containing BsgI and
                BamHI sites [5'gcccgcgcagcccgatccgcagaaaaag]
                [5'aattcttttttggatccggggtgcagc]"

ORIGIN
    Query Match      84.0%; Score 16.8; DB 13; Length 572;
    Best Local Similarity 90.0%; Pred. No. 6.4e+03;
    Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      1  GACGACGCTCTCCTTCCT 20
        |||||
Db      169 GACAACTCTCTCTTCCT 188

RESULT 14
BX249842
LOCUS      hr46d12.g1 WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum
            bicolor genomic clone hr46d12 5', genomic survey sequence.
DEFINITION
ACCESSION  BX249842
VERSION     BZ349842.1  GI:24912187
KEYWORDS   GSS.
SOURCE     Sorghum bicolor (sorghum)
ORGANISM   Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE  1 (bases 1 to 572)
AUTHORS   Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
            Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
            Zutavern, T., Palmer, L., McCombie, W.R. and Martienssen, R.A.
            Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
            Unpublished (2002)
            Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mcombie@cshl.org

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 28.353 Seconds  
(without alignments)  
661.956 Million cell updates/sec

Title: US-03-877-819B-33

Perfect score: 20  
Sequence: 1 gacagcgcctcccttct 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 180 summaries

Database : N Geneseq\_29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003as:\*

8: geneseqn2003bs:\*

9: geneseqn2003cs:\*

10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	20	6	AAL48167 Human HLA
C 2	20	100.0	20	6	AAL48199 Human HLA
C 3	20	100.0	20	6	AAL48197 Human HLA
C 4	20	100.0	20	6	AAL48183 Human HLA
C 5	16.8	84.0	48436	6	ABN89533 Human cor
C 6	16.8	84.0	160552	4	AAD02697 Human gly
C 7	16.4	82.0	546	4	AB122959 Drosophila
C 8	16.4	82.0	2546	4	AB122958 Drosophila
C 9	15.8	79.0	307	6	ABL98955 Omesa-con
C 10	15.8	79.0	850	6	ABQ49803 Oligonucle
C 11	15.8	79.0	850	6	ABQ49802 Oligonucle
C 12	15.8	79.0	1362	6	ABK72943 Bacillus
C 13	15.8	79.0	1686	5	AKS85172 DNA encod
C 14	15.8	79.0	2000	7	ADA73217 Rice gene
C 15	15.8	79.0	2000	7	ADA72949 Rice gene
C 16	15.8	79.0	3213	7	ABZ24255 Human SLC
C 17	15.8	79.0	3213	9	ADD18763 Human dis
C 18	15.8	79.0	3229	5	AA575476 DNA encod
C 19	15.8	79.0	3229	5	AA575478 DNA encod
C 20	15.8	79.0	6901	9	ADR09927 Novel DNA
C 21	15.8	79.0	7690	9	ABL33122 Human imm
C 22	15.8	79.0	14924	6	ABL32225 Human imm
C 23	15.8	79.0	14924	6	ABL54322 Chemical

C 24	15.8	79.0	177851	7	AAL57272	AAL57272	Abk157272	ba438B03-
C 25	15.4	77.0	496	8	ACL21232	ACL21232	Abk157272	DNA clone
C 26	15.4	77.0	525	8	ACL21237	ACL21237	Abk157272	DNA clone
C 27	15.4	77.0	556	8	ACL21234	ACL21234	Abk157272	DNA clone
C 28	15.4	77.0	562	8	ACL21236	ACL21236	Abk157272	DNA clone
C 29	15.4	77.0	574	8	ACL21228	ACL21228	Abk157272	DNA clone
C 30	15.4	77.0	610	6	ABQ15884	ABQ15884	Abk15884	Oligonucle
C 31	15.4	77.0	610	6	ABQ15885	ABQ15885	Abk15885	Oligonucle
C 32	15.4	77.0	673	6	ABQ42042	ABQ42042	Abq42042	Oligonucle
C 33	15.4	77.0	673	6	ABQ42043	ABQ42043	Abq42043	Oligonucle
C 34	15.4	77.0	713	8	ACL21235	ACL21235	Abk157272	DNA clone
C 35	15.4	77.0	716	8	ACL21233	ACL21233	Abk157272	DNA clone
C 36	15.4	77.0	852	6	ABN68486	ABN68486	Abn68486	Streptoco
C 37	15.4	77.0	1317	2	AA559113	AA559113	Aax59113	Mouse pan
C 38	15.4	77.0	1522	2	AA559115	AA559115	Aax59115	Mouse pan
C 39	15.4	77.0	2016	9	AD390631	AD390631	Adb90631	Guanine n
C 40	15.4	77.0	2402	9	AD390630	AD390630	Adb90630	Guanine n
C 41	15.4	77.0	5524	9	AD390625	AD390625	Adb90625	Mouse gua
C 42	15.4	77.0	34118	9	ADC86456	ADC86456	Adc86456	Human GPC
C 43	15.4	77.0	110000	7	AAL52246	AAL52246	Continuation (2 of	
C 44	15.4	77.0	227968	6	ABN83497	ABN83497	Abn83497	Human cdn
C 45	15.2	76.0	65	6	ABN27551	ABN27551	Abn27551	Rat splic
C 46	15.2	76.0	323	5	ABV48086	ABV48086	Abv48086	Human pro
C 47	15.2	76.0	459	8	ACH38985	ACH38985	Ach38985	Human foe
C 48	15.2	76.0	461	5	ABV18299	ABV18299	Abv18299	Human pro
C 49	15.2	76.0	488	8	ACH13744	ACH13744	Ach13744	Human adu
C 50	15.2	76.0	590	3	AAF14812	AAF14812	Aaf14812	Aspergill
C 51	15.2	76.0	590	7	ABZ54025	ABZ54025	Abz54025	Aspergill
C 52	15.2	76.0	776	6	ABQ48378	ABQ48378	Abq48378	Oligonucle
C 53	15.2	76.0	776	6	ABQ48379	ABQ48379	Abq48379	Oligonucle
C 54	15.2	76.0	780	6	ABQ49845	ABQ49845	Abq49845	Oligonucle
C 55	15.2	76.0	780	6	ABQ49844	ABQ49844	Abq49844	Oligonucle
C 56	15.2	76.0	949	4	AAH05625	AAH05625	Aah05625	Human cdn
C 57	15.2	76.0	1156	6	ABK63440	ABK63440	Abk63440	Rat seque
C 58	15.2	76.0	1156	9	ADB57852	ADB57852	Adb57852	Toxicity-
C 59	15.2	76.0	1156	9	ADB52358	ADB52358	Adb52358	Primary r
C 60	15.2	76.0	1156	9	ADB85151	ADB85151	Adb85151	Rat UDP-g
C 61	15.2	76.0	1212	6	ABQ69043	ABQ69043	Abq69043	Listeria
C 62	15.2	76.0	1453	3	AAH42323	AAH42323	Aah42323	Arabidops
C 63	15.2	76.0	1564	4	AAH14112	AAH14112	Aah14112	Human cdn
C 64	15.2	76.0	1916	4	AAH12684	AAH12684	Aah12684	Human cdn
C 65	15.2	76.0	1937	4	NAC91307	NAC91307	Nac91307	Human pol
C 66	15.2	76.0	1938	4	ABA08798	ABA08798	Abao8798	Human orp
C 67	15.2	76.0	2505	7	ABZ23069	ABZ23069	Abz23069	Human GPC
C 68	15.2	76.0	2700	5	AAH18502	AAH18502	Aah18502	Human cdn
C 69	15.2	76.0	2701	5	ABAJ5977	ABAJ5977	Abaj5977	Human ner
C 70	15.2	76.0	2934	4	ABL15887	ABL15887	Ab115887	Drosophila
C 71	15.2	76.0	3396	2	AAQ55515	AAQ55515	Aaq55515	PTase PT
C 72	15.2	76.0	3561	2	AAQ55970	AAQ55970	Aaq55970	PTase PT
C 73	15.2	76.0	3566	6	AB199237	AB199237	Abi99237	Mouse isc
C 74	15.2	76.0	4608	6	ABQ70898	ABQ70898	Abq70898	Listeria
C 75	15.2	76.0	4841	4	AAK52955	AAK52955	Aak52955	Human pol
C 76	15.2	76.0	4880	4	AAK51971	AAK51971	Aak51971	Human pol
C 77	15.2	76.0	4898	7	ABZ79896	ABZ79896	Abz79896	Human nuc
C 78	15.2	76.0	5204	5	ABL15886	ABL15886	Ab115886	Drosophila
C 79	15.2	76.0	12961	5	ABL20012	ABL20012	Aba20012	Human ner
C 80	15.2	76.0	18890	4	ABL13818	ABL13818	Ab113818	Drosophila
C 81	15.2	76.0	24606	4	AAK84723	AAK84723	Aak84723	Human imm
C 82	15.2	76.0	24606	4	AAK78367	AAK78367	Aak78367	Human imm
C 83	15.2	76.0	28564	9	ADE63609	ADE63609	Ad663609	Human gen
C 84	15.2	76.0	39887	4	AAK81263	AAK81263	Aak81263	Human imm
C 85	15.2	76.0	39887	4	AAK79153	AAK79153	Aak79153	Human imm
C 86	15.2	76.0	70419	7	AAQ56111	AAQ56111	Aaq56111	Human WNT
C 87	15.2	76.0	70419	9	ADA02473	ADA02473	Ada02473	Human WNT
C 88	15.2	76.0	70419	9	ADB72212	ADB72212	Adb72212	Human WNT
C 89	15.2	76.0	96583	3	AAF22297	AAF22297	Aaf22297	BAC conta
C 90	15.2	76.0	110000	6	ABX08336	ABX08336	Abx08336	Continuation (8 of
C 91	15.2	76.0	110000	6	ABQ67196	ABQ67196	Abq67196	Continuation (6 of
C 92	15.2	76.0	110000	6	ABQ69245	ABQ69245	Abq69245	Continuation (27 of
C 93	15.2	76.0	110000	6	ABA03041	ABA03041	Abao3041	Continuation (28 of
C 94	15	75.0	2063	9	ACF63557	ACF63557	Acf63557	Human cdn
C 95	15	75.0	2970	9	ACF79497	ACF79497	Acf79497	Cattle pl
C 96	15	75.0	24948	6	ABN85733	ABN85733	Abn85733	Mouse gen

97 14.8 74.0 235 4 AAH36766  
 c 98 14.8 74.0 271 7 ABX84599  
 99 14.8 74.0 284 7 ABX88350  
 100 14.8 74.0 345 7 ACF68272  
 101 14.8 74.0 524 6 ACF68272  
 c 102 14.8 74.0 525 6 ACF68272  
 c 103 14.8 74.0 564 9 ADD17229  
 c 104 14.8 74.0 564 9 ADD17229  
 c 105 14.8 74.0 597 5 AAK71493  
 c 106 14.8 74.0 621 6 AAK71493  
 c 107 14.8 74.0 642 6 AAK71493  
 c 108 14.8 74.0 642 6 AAK71493  
 c 109 14.8 74.0 732 2 AAK71493  
 c 110 14.8 74.0 739 4 AAK71493  
 c 111 14.8 74.0 801 6 AAK71493  
 c 112 14.8 74.0 801 6 AAK71493  
 c 113 14.8 74.0 890 4 AAK71493  
 c 114 14.8 74.0 911 6 AAK71493  
 c 115 14.8 74.0 911 6 AAK71493  
 c 116 14.8 74.0 1380 6 AAK71493  
 c 117 14.8 74.0 1754 7 AAK71493  
 c 118 14.8 74.0 1884 7 AAK71493  
 c 119 14.8 74.0 1937 7 AAK71493  
 c 120 14.8 74.0 2000 7 AAK71493  
 c 121 14.8 74.0 2039 5 AAK71493  
 c 122 14.8 74.0 2318 4 AAK71493  
 c 123 14.8 74.0 2803 7 AAK71493  
 c 124 14.8 74.0 2821 9 AAK71493  
 c 125 14.8 74.0 2966 3 AAK71493  
 c 126 14.8 74.0 3361 4 AAK71493  
 c 127 14.8 74.0 3475 5 AAK71493  
 c 128 14.8 74.0 3480 3 AAK71493  
 c 129 14.8 74.0 3936 4 AAK71493  
 c 130 14.8 74.0 3950 3 AAK71493  
 c 131 14.8 74.0 4070 4 AAK71493  
 c 132 14.8 74.0 4070 4 AAK71493  
 c 133 14.8 74.0 4070 4 AAK71493  
 c 134 14.8 74.0 4070 4 AAK71493  
 c 135 14.8 74.0 4380 6 AAK71493  
 c 136 14.8 74.0 4710 4 AAK71493  
 c 137 14.8 74.0 5243 6 AAK71493  
 c 138 14.8 74.0 5243 6 AAK71493  
 c 139 14.8 74.0 5401 4 AAK71493  
 c 140 14.8 74.0 5401 4 AAK71493  
 c 141 14.8 74.0 6169 4 AAK71493  
 c 142 14.8 74.0 7028 4 AAK71493  
 c 143 14.8 74.0 7694 4 AAK71493  
 c 144 14.8 74.0 7694 4 AAK71493  
 c 145 14.8 74.0 8088 4 AAK71493  
 c 146 14.8 74.0 8535 2 AAK71493  
 c 147 14.8 74.0 9160 4 AAK71493  
 c 148 14.8 74.0 10069 6 AAK71493  
 c 149 14.8 74.0 10213 9 AAK71493  
 c 150 14.8 74.0 13310 4 AAK71493  
 c 151 14.8 74.0 13310 4 AAK71493  
 c 152 14.8 74.0 17481 4 AAK71493  
 c 153 14.8 74.0 17481 4 AAK71493  
 c 154 14.8 74.0 17481 4 AAK71493  
 c 155 14.8 74.0 33239 8 AAK71493  
 c 156 14.8 74.0 42334 9 AAK71493  
 c 157 14.8 74.0 42334 9 AAK71493  
 c 158 14.8 74.0 89328 7 AAK71493  
 c 159 14.8 74.0 110000 7 AAK71493  
 c 160 14.8 74.0 110000 7 AAK71493  
 c 161 14.8 74.0 110000 7 AAK71493  
 c 162 14.8 74.0 110000 7 AAK71493  
 c 163 14.8 74.0 110000 7 AAK71493  
 c 164 14.8 74.0 302250 6 AAK71493  
 c 165 14.8 74.0 349980 6 AAK71493  
 c 166 14.8 74.0 349980 6 AAK71493  
 c 167 14.4 72.0 330 6 AAK71493  
 c 168 14.4 72.0 405 4 AAK71493  
 c 169 14.4 72.0 412 6 AAK71493

## ALIGNMENTS

RESULT 1  
 AAL48167/c  
 ID AAL48167 standard; DNA; 20 BP.  
 XX AC AAL48167;  
 XX DT 01-OCT-2002 (first entry)  
 XX DE Human HLA DP31 locus polymorphism address tag sequence #3.  
 XX DE DE  
 KW KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
 KW KW flow cytometry; human; DP31; address tag; ss.  
 OS Homo sapiens.  
 XX WO200194639-A1.  
 PD 13-DEC-2001.  
 XX 07-JUN-2001; 2001WO-US018590.  
 PR 08-JUN-2000; 2000US-0210759P.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX White PS, Torney DC;  
 FI WPI; 2002-566450/60.  
 DR DR  
 XX XX Identifying sequences useful as address/capture tags for flow cytometry  
 PT based minisequencing, by generating tag sequences and rejecting sequences  
 PT based on certain parameters e.g. sequences which form stable hairpins.  
 XX XX Disclosure; Page 9; 35pp; English.  
 PS The present invention relates to a method of identifying sequences useful  
 CC as addresses/capture tags, involving rejecting sequences having common sub-  
 CC sequences with a sub-sequence length greater than specified number of  
 CC bases, and sequences which can form stable hairpins and stable dimers  
 CC from a sample of oligonucleotides, and selecting those sequences in the  
 CC sample that would hybridise to their respective complements with a high  
 CC degree of specificity. The method is useful for identifying a set of  
 CC sequences useful as address/capture tags which can be used for  
 CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
 CC cytometry assay. The present sequence is an address tag described in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 20 BP; 6 A; 3 C; 9 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred.No. 6.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGAGGCTCCCTTCTCT 20  
 DB 20 GACGAGGCTCCCTTCTCT 1



```

RESULT 2
AAL48199/c
ID AAL48199 standard; DNA; 20 BP.
XX
XX AAL48199;
XX
XX 01-OCT-2002 (first entry)
XX
XX Human HLA DPB1 locus polymorphism address tag sequence #1.
XX
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
KW flow cytometry; human; DPB1; address tag; ss.
XX
XX Homo sapiens.
XX
XX WO200194639-A1.
XX
XX 13-DEC-2001.
XX
XX 07-JUN-2001; 2001WO-US018590.
XX
XX 08-JUN-2000; 2000US-0210759P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX White PS, Torney DC;
XX
XX WPI; 2002-566450/60.
XX
XX Identifying sequences useful as address/capture tags for flow cytometry
PT based minisequencing, by generating tag sequences and rejecting sequences
PT based on certain parameters e.g. sequences which form stable hairpins.
XX
XX Disclosure; Page 14; 35pp; English.
XX
XX The present invention relates to a method of identifying sequences useful
CC as address/capture tags, involving rejecting sequences having common sub-
CC sequences with a sub-sequence length greater than specified number of
CC bases, and sequences which can form stable hairpins and stable dimers
CC from a sample of oligonucleotides, and selecting those sequences in the
CC sample that would hybridise to their respective complements with a high
CC degree of specificity. The method is useful for identifying a set of
CC sequences useful as address/capture tags which can be used for
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
CC cytometry assay. The present sequence is an address tag described in the
CC exemplification of the invention
XX
XX Sequence 20 BP; 6 A; 3 C; 9 G; 2 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 6; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 6.4;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GACGAGCCTCTCCTTTCT 20
XX |||||
XX Db 20 GACGAGCCTCTCCTTTCT 1
XX
XX RESULT 3
XX AAL48197
XX ID AAL48197 standard; DNA; 20 BP.
XX
XX AAL48197;
XX
XX 01-OCT-2002 (first entry)
XX
XX Human HLA DPB1 locus polymorphism multiplex capture sequence #1.
XX
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
KW flow cytometry; human; DPB1; capture tag; ss.
XX
XX Homo sapiens.
XX

```

```

XX
XX WO200194639-A1.
XX
XX 13-DEC-2001.
XX
XX 07-JUN-2001; 2001WO-US018590.
XX
XX 08-JUN-2000; 2000US-0210759P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX White PS, Torney DC;
XX
XX WPI; 2002-566450/60.
XX
XX Identifying sequences useful as address/capture tags for flow cytometry
PT based minisequencing, by generating tag sequences and rejecting sequences
PT based on certain parameters e.g. sequences which form stable hairpins.
XX
XX Disclosure; Page 14; 35pp; English.
XX
XX The present invention relates to a method of identifying sequences useful
CC as address/capture tags, involving rejecting sequences having common sub-
CC sequences with a sub-sequence length greater than specified number of
CC bases, and sequences which can form stable hairpins and stable dimers
CC from a sample of oligonucleotides, and selecting those sequences in the
CC sample that would hybridise to their respective complements with a high
CC degree of specificity. The method is useful for identifying a set of
CC sequences useful as address/capture tags which can be used for
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
CC cytometry assay. The present sequence is a capture tag described in the
CC exemplification of the invention
XX
XX Sequence 20 BP; 2 A; 9 C; 3 G; 6 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 6; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 6.4;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GACGAGCCTCTCCTTTCT 20
XX |||||
XX Db 1 GACGAGCCTCTCCTTTCT 20
XX
XX RESULT 4
XX AAL48183
XX ID AAL48183 standard; DNA; 20 BP.
XX
XX AAL48183;
XX
XX 01-OCT-2002 (first entry)
XX
XX Human HLA DPB1 locus polymorphism multiplex capture sequence #3.
XX
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
KW flow cytometry; human; DPB1; address tag; ss.
XX
XX Homo sapiens.
XX
XX WO200194639-A1.
XX
XX 13-DEC-2001.
XX
XX 07-JUN-2001; 2001WO-US018590.
XX
XX 08-JUN-2000; 2000US-0210759P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX White PS, Torney DC;
XX
XX WPI; 2002-566450/60.
XX

```

PT Identifying sequences useful as address/capture tags for flow cytometry  
 PT based on minisequencing, by generating tag sequences and rejecting sequences  
 PT based on certain parameters e.g. sequences which form stable hairpins.

XX Disclosure; Page 9; 35pp; English.

XX The present invention relates to a method of identifying sequences useful  
 CC as address/capture tags, involving rejecting sequences having common sub-  
 CC sequences with a sub-sequence length greater than specified number of  
 CC bases, and sequences which can form stable hairpins and stable dimers  
 CC from a sample of oligonucleotides, and selecting those sequences in the  
 CC sample that would hybridise to their respective complements with a high  
 CC degree of specificity. The method is useful for identifying a set of  
 CC sequences useful as address/capture tags which can be used for  
 CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
 CC cytometry assay. The present sequence is an address tag described in the  
 CC exemplification of the invention

XX Sequence 20 BP; 2 A; 9 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGACGCTCTCCTTTCT 20  
 Db 1 GACGACGCTCTCCTTTCT 20

RESULT 5

ABN89533  
 ID ABN89533 standard; DNA; 48436 BP.

XX AC ABN89533;

XX 05-SEP-2002 (first entry)

XX Human corneal N-acetylglucosamine-6-sulfotransferase DNA SEQ ID NO:38.

XX Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;  
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;  
 KW ophthalmological; chromosome 16q22; gene; ds.

XX Homo sapiens.

XX US2002061562-A1.

XX 23-MAY-2002.

XX 09-AUG-2001; 2001US-0927602.

XX 11-AUG-2000; 2000US-00638211.

PR 11-AUG-2000; 2000US-0325773P.

XX (FUKU/) FUKUDA M N.

XX (AKAM/) AKAMA T O.

XX Fukuda MN, Akama TO;

XX WPI; 2002-507643/54.

XX New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase,  
 PT useful for treatment, monitoring and diagnosis of macular corneal  
 PT dystrophy.

XX Claim 35; Page 31-53; 69pp; English.

XX The present sequence encodes human corneal N-acetylglucosamine-6-  
 CC sulfotransferase (GlcNAc6ST) (I), which is able to catalyse sulfation of  
 CC keratan sulfate (KS). Also described is a method for monitoring the  
 CC effect of treatments for macular corneal dystrophy (MCD), and detecting  
 CC susceptibility to MCD. (I) is located to chromosome 16q22, and has  
 CC ophthalmological activity. (I) can be used to treat or prevent macular

CC corneal dystrophy types I or II. (I) makes possible treatment of MCD  
 CC without requiring keratinoplasty or keratectomy

XX Sequence 48436 BP; 11653 A; 11904 C; 11645 G; 13234 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 6; Length 48436;

Best Local Similarity 90.0%; Pred. No. 2.8e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACGACGCTCTCCTTTCT 20

Db 2118 GATGACGCTCTCCTTTCT 2-37

RESULT 6

AAD02697

ID AAD02697 standard; DNA; 160552 BP.

XX AC AAD02697;

XX 02-MAY-2001 (first entry)

XX Human glycosyl sulfotransferase-4 (GST-4) genomic DNA.

XX Human; glycosyl sulfotransferase-4; GST-4; immunosuppressive; therapy;  
 KW selectin binding inhibitor; gene therapy; inflammation;  
 KW systemic lupus erythematosus; SLB; rheumatoid arthritis; diabetes;  
 KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;  
 KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adenitis;  
 KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;  
 KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;  
 KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;  
 KW asthma; hypersensitivity; rheumatic fever; tissue rejection;  
 KW chromosome 15q23.1; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT exon 32847..32922

FT /tag= a

FT /number= 1

FT /label= 4a 5U4

FT 32923..35592

FT /tag= b

FT /cons\_splice= (5'site:NO, 3'site:YES)

FT 35593..35674

FT /tag= c

FT /number= 2

FT /label= 4a 5U3

FT 35675..45093

FT /tag= d

FT 45094..45185

FT /tag= e

FT /number= 3

FT /label= 4a 5U2

FT 45186..46633

FT /tag= f

FT /cons\_splice= (5'site:NO, 3'site:NO)

FT 46634..46700

FT /tag= g

FT /number= 4

FT /label= 4a 5U1

FT 46701..47938

FT /tag= h

FT /cons\_splice= (5'site:YES, 3'site:NO)

FT 47939..49746

FT /tag= i

FT /number= 5

FT /note= "Includes 17 base pairs of 5'UTR, the ORF and all

FT of 3'UTR"

FT 5'UTR

FT 47939..47955

FT /tag= j

FT /note= "Portion of 5' untranslated region (5'UTR)"

FT CDS 47956. .49128  
 FT /tag= k  
 FT /product= "Human glycosyl transferase-4alpha (GST-4alpha)"  
 FT 3'UTR 49129. .49746  
 FT /tag= l  
 FT exon 83257. .83347  
 FT /tag= m  
 FT /label= 4a.5U2  
 FT intron 83348. .96412  
 FT /tag= n  
 FT /cons\_splice= (5'site:NO, 3'site:NO)  
 FT exon 96413. .96484  
 FT /tag= o  
 FT /label= 4a.5U1  
 FT intron 96485. .98456  
 FT /tag= p  
 FT /cons\_splice= (5'site:NO, 3'site:NO)  
 FT exon 98457. .99968  
 FT /tag= q  
 FT /note= "Includes 17 base pairs of 5'UTR, the ORF and all of 3'UTR"  
 FT 5'UTR 98457. .98473  
 FT /tag= r  
 FT /note= "Portion of 5' untranslated region (5'UTR)"  
 FT CDS 98474. .99661  
 FT /tag= s  
 FT /product= "Human glycosyl transferase-4beta (GST-4beta)"  
 FT 3'UTR 99662. .99968  
 FT /tag= t  
 FT WO200106015-A1.  
 FT PN  
 FT 25-JAN-2001.  
 FT PD  
 FT 19-JUL-2000; 2000WO-US019741.  
 FT XX  
 FT 20-JUL-1999; 99US-0144694P.  
 FT PR 13-JUN-2000; 2000US-00593828.  
 FT XX  
 FT PA (REGC ) UNIV CALIFORNIA.  
 FT XX  
 FT Rosen SD, Lee JK, Hemmerich S;  
 FT PI WPI; 2001-138471/14.  
 FT DR P-PSDB; AAY72639, AAY72640.  
 FT XX  
 FT New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for  
 FT diagnostic and therapeutic agent screening applications.  
 FT XX  
 FT Example 1; Page 62-104; 128pp; English.  
 FT XX  
 FT The present sequence is human glycosyl sulfotransferase-4 (GST-4) genomic  
 CC DNA encoding GST-4alpha and GST-4beta. GST-4 gene is found on chromosome  
 CC 16q23.1. GST is a type 2 membrane protein useful for inhibiting a binding  
 CC event between a selectin and a selectin ligand, which comprises  
 CC contacting the selectin with a non-sulphated selectin ligand. GST and a  
 CC small molecular agent that inhibits the sulphation activity of GST. GST  
 CC is also useful in inhibiting a selectin mediated binding event. GST is  
 CC useful in gene therapy to treat disorders such as acute or chronic  
 CC inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis,  
 CC polyarthritis nodosa, polymyositis, dermatomyositis, systemic sclerosis,  
 CC diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome,  
 CC Hashimoto's disease, Grave's disease, adrenalitis, hypoparathyroidism,  
 CC pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative  
 CC colitis, dermatitis, myocarditis, regional enteritis, adult respiratory  
 CC distress syndrome, infantile eczema, psoriasis lichen planus, allergic  
 CC rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue  
 CC rejection during transplantation  
 FT SQ Sequence 160552 BP; 40281 A; 37573 C; 38015 G; 44564 T; 0 U; 119 Other;  
 Query Match 84.0%; Score 16.8; DB 4; Length 160552;

Best Local Similarity 90.0%; Pred. No. 2.9e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GACGACGCTCTCCTTCCT 20  
 |||||  
 DB 53027 GATGACGCTCTGCTTCCT 53046  
 RESULT 7  
 ABL22959/C  
 ID ABL22959 standard; DNA; 546 BP.  
 XX  
 AC ABL22959;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20350.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 20350; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signaling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABBS72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 546 BP; 113 A; 193 C; 167 G; 73 T; 0 U; 0 Other;  
 Query Match 82.0%; Score 16.4; DB 4; Length 546;  
 Best Local Similarity 94.4%; Pred. No. 3.6e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GACGACGCTCTCCTTCCT 18  
 |||||  
 DB 466 GACGACGCTCTCCTTGC 449  
 RESULT 8  
 ABL22958/C  
 ID ABL22958 standard; DNA; 2546 BP.  
 XX  
 AC ABL22958;  
 XX  
 DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20347.  
 XX XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US009231.  
 XX PR 23-MAR-2000; 2000US-0191637P.  
 XX PR 11-JUL-2000; 2000US-00614150.  
 XX PA (PEKE ) PE CORP NY.  
 XX PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX DR WPI; 2001-656860/75.  
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX PS Claim 1; SEQ ID NO 20347; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16175-ABL30511), expressed DNA  
 CC sequences (AB01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABBS72072). The sequence data for this parent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 2546 BP; 753 A; 609 C; 568 G; 616 T; 0 U; 0 Other;  
 Query Match 82.0%; Score 16.4; DB 4; Length 2546;  
 BestLocal Similarity 94.4%; Pred. No. 3.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GACGACGCTCTCCTTTC 18  
 Db 1466 GACGACGCTCTCCTTGC 1449  
 RESULT 9  
 ABL98955/c  
 ID ABL98955 standard; DNA; 307 BP.  
 XX AC ABL98955;  
 XX 12-JUL-2002 (first entry)  
 DE Omega-conopeptide Ra6.4 encoding DNA.  
 XX Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;  
 KW neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;  
 KW antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;  
 KW anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;  
 KW neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;  
 KW stroke; cerebrovascular accident; brain trauma; spinal chord trauma;  
 KW drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;  
 KW migraine; inflammation; cardiovascular disorder; psychiatric disorder;  
 KW psychosis; anxiety; schizophrenia; ds.  
 XX OS Conus rattus.  
 XX DN WO200218632-A2.  
 XX

PN WO200207675-A2.  
 XX 31-JAN-2002.  
 XX PF 23-JUL-2001; 2001WO-US023041.  
 XX PR 21-JUL-2000; 2000US-0219616P.  
 XX PR 05-FEB-2001; 2001US-0265888P.  
 XX PA (UTAH ) UNIV UTAH RES FOUND.  
 XX PA (COGN-) COGNETIX INC.  
 XX PI Olivera BM, McIntosh JM, Watkins M, Garrett JE, Shon K;  
 XX PI Jacobsen R, Jones RM, Cartier GE;  
 XX DR WPI; 2002-257318/30.  
 XX DR P-PSDB; ABB96696.  
 XX PT New omega-conopeptides useful for treating disorders associated with  
 PT voltage gated ion channels e.g. pain, inflammation, neurologic or  
 PT cardiovascular disorders.  
 XX PS Claim 3; Page 70; 195pp; English.  
 CC The invention relates to isolated omega-conopeptides, nucleic acid  
 CC sequences encoding them, and propeptide sequences. The activity of the  
 CC peptides of the invention may be described as, analgesic, anticonvulsant,  
 CC vasotropic, cardiant, neuroprotective, cerebroprotective, cardiovascular,  
 CC antiinflammatory, antimigraine, antidiabetic, tranquiliser, vulnerary,  
 CC anxiolytic, neuroleptic and neuroleptic. Peptides of the invention act  
 CC by modulating the activity of voltage gated ion channels. They may be  
 CC used for treating or preventing disorders associated with voltage gated  
 CC ion channels such as neurological disorders, e.g. seizure (associated  
 CC with epilepsy), neurotoxic injury associated with conditions of hypoxia,  
 CC anoxia, ischaemia, stroke, cerebrovascular accident, brain or spinal  
 CC chord trauma, drowning, suffocation, perinatal asphyxia or hypoglycaemic  
 CC events; pain e.g. migraine; inflammation or cardiovascular disorders.  
 CC They may also be used for treating psychiatric disorders e.g. psychosis,  
 CC anxiety or schizophrenia. The analgesic agents of the invention show  
 CC diminished side effects and toxicity, and are non-addictive. The  
 CC sequences given in records ABL98855-ABL98956 represent omega-conopeptide  
 CC encoding DNA's  
 XX SQ Sequence 307 BP; 58 A; 84 C; 82 G; 83 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.8; DB 6; Length 307;  
 BestLocal Similarity 89.5%; Pred. No. 6.7e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 ACGACGCTCTCCTTTCCT 20  
 Db 105 ACGACGCTCTCCTTACCT 87  
 RESULT 10  
 ABL98803  
 ID ABL98803 standard; DNA; 850 BP.  
 XX AC ABL98803;  
 XX 12-JUL-2002 (first entry)  
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 36394.  
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX OS Homo sapiens.  
 XX DN WO200218632-A2.  
 XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 694.871 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819b-34

Perfect score: 23

Sequence: 1 cggaccatgtgcaacttatgcc 23

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_man.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	23	100.0	26	BD104339	BD104339 Kit and m
2	23	100.0	246	AF346471	AF346471 Homo sapi
3	23	100.0	246	HS0103X2	X82390 H.sapiens M
4	23	100.0	246	HS0201X2	X82394 H.sapiens M
5	23	100.0	246	HS1ADPA1	X78198 H.sapiens H
6	23	100.0	252	HS248473	Z48473 H.sapiens H
7	23	100.0	257	AF165160	AF165160 Homo sapi
8	23	100.0	258	AF118120	AF118120 Homo sapi
9	23	100.0	267	AX237167	AX237167 Sequence
10	23	100.0	267	AX237352	AX237352 Sequence
11	23	100.0	268	HS1ADPAX	X83610 H.sapiens H
12	23	100.0	272	AX237066	AX237066 Sequence
13	23	100.0	286	HSU87556	U87556 Homo sapien
14	23	100.0	287	AF015295	AF015295 Homo sapi
15	23	100.0	294	AX237304	AX237304 Sequence
16	23	100.0	294	AX237554	AX237554 Sequence
17	23	100.0	326	AF013767	AF013767 Homo sapi
18	23	100.0	396	BD058253	BD058253 Secreted
19	23	100.0	466	BD058339	BD058339 Secreted
20	23	100.0	476	AX884252	AX884252 Sequence
21	23	100.0	476	BD023862	BD023862 Sequence
22	23	100.0	576	AX884251	AX884251 Sequence
23	23	100.0	576	BD023861	BD023861 Sequence
24	23	100.0	661	E00485	E00485 DNA sequence
25	23	100.0	661	I03086	I03086 Sequence 5
26	23	100.0	690	S40633	S40633 HLA class I
27	23	100.0	818	HUMHDPDA	X00457 Human mRNA
28	23	100.0	1048	HSSBA1	X03088 Sequence 7
29	23	100.0	1140	E00484	E00484 DNA sequence
30	23	100.0	1201	E00484	E00484 DNA sequence
31	23	100.0	1201	I03006	I03006 Sequence 6
32	23	100.0	1201	I03423	I03423 Sequence 6
33	23	100.0	1201	BC009956	BC009956 Homo sapi
34	23	100.0	1259	AX552229	AX552229 Sequence
35	23	100.0	1480	AX780118	AX780118 Sequence
36	23	100.0	1480	AX780119	AX780119 Sequence
37	23	100.0	2986	HUMHEDRC02	X23984 Human MHC C
38	23	100.0	14646	HS1ASBA	X03100 Human HLA-S
39	23	100.0	64380	EX120009	EX120009 Human DNA
40	23	100.0	106728	AL805913	AL805913 Human DNA
41	23	100.0	124899	AL645931	AL645931 Human DNA
42	23	100.0	181228	AC011086	AC011086 Homo sapi
43	23	100.0	187964	AL662824	AL662824 Human DNA
44	22	95.7	214	HUMHEDRC1AX	LI1641 Human MHC C
45	21	91.3	244	AF098794	AF098794 Homo sapi
46	20	87.0	21	AR095334	AR095334 Sequence
47	20	87.0	21	AR288212	AR288212 Sequence
48	20	87.0	21	AR427863	AR427863 Sequence
49	18.8	81.7	1834	HSSBA2P	X01634 Human HLA-S
50	18.8	81.7	6565	HUMHESXA	MI1591 Human MHC C
51	18.8	81.7	22847	HS00826Q1	AL049813 Human DNA
52	18.8	81.7	93842	AL845446	AL845446 Human DNA
53	18.8	81.7	158033	AL645940	AL645940 Human DNA
54	18.8	81.7	175737	HS1033B10	AL031228 Human DNA
55	18.8	81.7	201923	AC121772	AC121772 Mus muscu
56	18.8	81.7	239439	AC117573	AC117573 Mus muscu
57	18.4	80.0	150206	AC125780	AC125780 Oryza sat
58	18.2	79.1	830	BV062317	BV062317 S212P6441
59	18.2	79.1	151087	BX649549	BX649549 Mouse DNA
60	18.2	79.1	161700	AC125261	AC125261 Mus muscu
61	18.2	79.1	182725	AC102786	AC102786 Mus muscu
62	18.2	79.1	277309	AC102788	AC102788 Mus muscu
63	18	78.3	279	AF076284	AF076284 Homo sapi
64	18	78.3	279	AF076285	AF076285 Homo sapi
65	18	78.3	51409	AF024666	AF024666 Corynebact

```

c 66 18 78.3 157638 9 AC064833 Homo sapi
c 67 18 78.3 188317 2 AC093837 Homo sapi
c 68 17.8 77.4 2142 3 AF519419 Strongylo
c 69 17.8 77.4 86061 9 AL591472 Human DNA
c 70 17.8 77.4 82656 8 AP006087 Iotus cor
c 71 17.8 77.4 140817 10 AC127231 Mus muscu
c 72 17.8 77.4 178588 2 AC114234 Rattus no
c 73 17.8 77.4 202316 10 AC138134 Mus muscu
c 74 17.8 77.4 208524 9 AC061709 Homo sapi
c 75 17.8 77.4 217227 2 AC111530 Rattus no
c 76 17.8 77.4 229655 2 AC096929 Rattus no
c 77 17.8 77.4 244323 2 AC114128 Rattus no
c 78 17.8 77.4 270251 2 BX649488 Danio rer
c 79 17.4 75.7 11882 8 AY245981 Aegilops
c 80 17.4 75.7 143654 8 AC080019 Human DNA
c 81 17.4 75.7 163770 2 AF005648 Oryza sat
c 82 17.4 75.7 174877 2 AP004818 Oryza sat
c 83 17.4 75.7 195658 2 AC125196 Mus muscu
c 84 17.4 75.7 216052 10 AC111036 Mus muscu
c 85 17.4 75.7 304297 8 AC017088 Oryza sat
c 86 17.2 74.8 409 6 AC069638 Sequence
c 87 17.2 74.8 429 6 AX070029 Sequence
c 88 17.2 74.8 432 6 AX071881 Sequence
c 89 17.2 74.8 790 11 BV019301 Sequence
c 90 17.2 74.8 1105 9 BC047087 Homo sapi
c 91 17.2 74.8 1143 8 AY277395 Eumymus
c 92 17.2 74.8 1263 10 AB084158 Cavia por
c 93 17.2 74.8 1752 9 HSM803219 Homo sapi
c 94 17.2 74.8 1775 9 AK058121 Homo sapi
c 95 17.2 74.8 1933 8 X66608 B.napus mRN
c 96 17.2 74.8 2270 9 BC63521 Homo sapi
c 97 17.2 74.8 3220 2 BC001196 Homo sapi
c 98 17.2 74.8 44534 2 AC101338 Homo sapi
c 99 17.2 74.8 78016 2 AC084173 Homo sapi
c 100 17.2 74.8 110000 2 AC130479 Homo sapi
c 101 17.2 74.8 118648 9 AL138742 Human DNA
c 102 17.2 74.8 128959 2 AC006384 Homo sapi
c 103 17.2 74.8 131095 2 AL590071 Homo sapi
c 104 17.2 74.8 145673 2 AL391068 Homo sapi
c 105 17.2 74.8 151865 2 AC129931 Mus muscu
c 106 17.2 74.8 161532 9 AC137255 Rattus no
c 107 17.2 74.8 168190 2 AC101775 Mus muscu
c 108 17.2 74.8 170149 2 AC137255 Rattus no
c 109 17.2 74.8 171537 9 AC006972 Homo sapi
c 110 17.2 74.8 176075 9 AC017079 Homo sapi
c 111 17.2 74.8 177274 9 AP000424 Homo sapi
c 112 17.2 74.8 177433 2 AC142248 Mus muscu
c 113 17.2 74.8 178988 2 AC107687 Mus muscu
c 114 17.2 74.8 184007 9 AC116099 Homo sapi
c 115 17.2 74.8 185807 2 AC113112 Mus muscu
c 116 17.2 74.8 188297 2 AC131035 Mus muscu
c 117 17.2 74.8 198297 2 AC025655 Homo sapi
c 118 17.2 74.8 201387 2 AC025655 Homo sapi
c 119 17.2 74.8 206231 9 AL592309 Human DNA
c 120 17.2 74.8 206236 2 AC137152 Mus muscu
c 121 17.2 74.8 208276 2 AC102901 Mus muscu
c 122 17.2 74.8 212862 2 AC097294 Rattus no
c 123 17.2 74.8 216155 2 AC130503 Rattus no
c 124 17.2 74.8 216328 2 AC102247 Mus muscu
c 125 17.2 74.8 216615 2 AC128550 Rattus no
c 126 17.2 74.8 221054 2 AC009882 Homo sapi
c 127 17.2 74.8 234662 2 AC112059 Rattus no
c 128 17.2 74.8 238195 2 AC098115 Rattus no
c 129 17.2 74.8 239312 2 AC111457 Rattus no
c 130 17.2 74.8 241679 2 AC020861 Mus muscu
c 131 17.2 74.8 244328 2 AC094938 Rattus no
c 132 17.2 74.8 246366 2 AC094648 Rattus no
c 133 17.2 74.8 260052 2 AC094071 Rattus no
c 134 17.2 74.8 270097 2 AC110901 Mus muscu
c 135 17.2 74.8 287664 2 AC109210 Mus muscu
c 136 17.2 74.8 288605 2 AC101820 Mus muscu
c 137 17.2 74.8 294494 2 AC098385 Rattus no
c 138 17.2 74.8 298050 1 AL935260 Lactobaci

```

## ALIGNMENTS

## RESULT 1

LOCUS BD104339 26 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Kit and method for determining HLA type.

ACCESSION BD104339

VERSION BD104339.1 GI:22649913

KEYWORDS WO 0192572-A/443.

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 26)

AUTHORS Inoko,H., Kagiya,T., Ichihara,T., Matsumura,Y., Moriya,S. and Nishida,M.

TITLE Kit and method for determining HLA type

JOURNAL Patent: WO 0192572-A 443 06-DEC-2001;

NISSHIKO INDUSTRIES INC,SYSTEM RESEARCH INC,HIDETOSHI INOKO, TAEKO

KAGIYA, TATSUO ICHIHARA,YOSHIYUKI MATSUMURA,SHOGO MORIYA,MICHIO

NISHIDA

COMMENT OS Artificial Sequence

PN WO 0192572-A/443

PD 06-DEC-2001

PF 01-JUN-2001 WO 2001JP004662

PR 01-JUN-2000 JP 00P 164798

PI HIDETOSHI INOKO,TAEKO KAGIYA,TATSUO ICHIHARA,YOSHIYUKI PI

MATSUMURA,SHOGO MORIYA,MICHIO NISHIDA

PI SHOGO MORIYA,MICHIO NISHIDA

```

139 17.2 74.8 300853 2 AC135145 Rattus no
140 17 73.9 265 9 AF092049 Homo sapi
c 141 17 73.9 102816 2 BX296529 Homo sapi
142 17 73.9 105836 9 AC068780 Homo sapi
c 143 17 73.9 113301 2 AC138331 Homo sapi
c 144 17 73.9 119943 9 AC144940 Mus muscu
145 17 73.9 146527 10 BX088584 Mouse DNA
c 146 17 73.9 166197 2 AC073552 Homo sapi
c 147 17 73.9 248162 2 AC130584 Rattus no
148 17 73.9 265566 2 AC123187 Rattus no
149 17 73.9 286347 2 AC106098 Rattus no
150 16.8 73.0 1677 3 AY336497 Schistos
c 151 16.8 73.0 4043 8 AF040083 Sclerotin
c 152 16.8 73.0 4043 8 AF342302 Sclerotin
c 153 16.8 73.0 4043 8 AF342303 Sclerotin
c 154 16.8 73.0 4043 8 AF342304 Sclerotin
c 155 16.8 73.0 4043 8 AF342305 Sclerotin
c 156 16.8 73.0 4043 8 AF342306 Sclerotin
c 157 16.8 73.0 4044 8 AF040081 Sclerotin
c 158 16.8 73.0 4044 8 AF342010 Sclerotin
c 159 16.8 73.0 4044 8 AF342011 Sclerotin
c 160 16.8 73.0 4044 8 AF342012 Sclerotin
c 161 16.8 73.0 4044 8 AF342013 Sclerotin
c 162 16.8 73.0 4044 8 AF342014 Sclerotin
c 163 16.8 73.0 4044 8 AF342015 Sclerotin
c 164 16.8 73.0 4044 8 AF342016 Sclerotin
c 165 16.8 73.0 4044 8 AF342017 Sclerotin
c 166 16.8 73.0 4044 8 AF342018 Sclerotin
c 167 16.8 73.0 4044 8 AF342019 Sclerotin
c 168 16.8 73.0 4044 8 AF342020 Sclerotin
c 169 16.8 73.0 4044 8 AF342021 Sclerotin
c 170 16.8 73.0 4044 8 AF342022 Sclerotin
c 171 16.8 73.0 4044 8 AF342023 Sclerotin
c 172 16.8 73.0 4044 8 AF342024 Sclerotin
c 173 16.8 73.0 4044 8 AF342025 Sclerotin
c 174 16.8 73.0 4044 8 AF342026 Sclerotin
c 175 16.8 73.0 4044 8 AF342027 Sclerotin
c 176 16.8 73.0 4044 8 AF342028 Sclerotin
c 177 16.8 73.0 4044 8 AF342029 Sclerotin
c 178 16.8 73.0 4044 8 AF342030 Sclerotin
c 179 16.8 73.0 4044 8 AF342031 Sclerotin
c 180 16.8 73.0 4044 8 AF342032 Sclerotin

```

PC C12M1/00.C12M15/09.G01N33/53  
 CC Description of Artificial Sequence:primer  
 FH Key Location/Qualifiers  
 FT source 1..26  
 FT Location/Qualifiers  
 FEATURES  
 source 1..26  
 /organism="Artificial Sequence".  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

ORIGIN  
 Query Match 100.0%; Score 23; DB 6; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 0.092;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 4 CGGACCATGTGTCAACTTATGCC 26

RESULT 2  
 AF346471  
 LOCUS 246 bp DNA linear PRI 27-MAR-2001  
 DEFINITION Homo sapiens MHC class II antigen (HLA-DPA1) gene, HLA-DPA1-new  
 allele, partial cds.  
 ACCESSION AF346471  
 VERSION AF346471.1 GI:13448661

KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 246)  
 Grams,S.E., Begovich,A. and Mangaccat,J.  
 One new DPA1 Allele  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 246)  
 Grams,S.E., Begovich,A. and Mangaccat,J.  
 Direct Submission  
 TITLE Submitted (07-FEB-2001) Human Genetics, Roche Molecular Systems,  
 1145 Atlantic Ave., Alameda, CA 94501, USA  
 JOURNAL Location/Qualifiers

FEATURES  
 source 1..246  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="6p"  
 <1..>246  
 /gene="HLA-DPA1"  
 /allele="new"  
 <1..>246  
 /gene="HLA-DPA1"  
 /product="MHC class II antigen"  
 <1..>246  
 /gene="HLA-DPA1"  
 /codon\_start=3  
 /product="MHC class II antigen"  
 /protein\_id="AAK27152.1"  
 /db\_xref="GI:13448662"  
 /translation="DHYSTYAAFVQTHREPTGEERFDDDDMFYVLDLKKETVHLEEFGRAPFPAQGLAIALLNNLNLQRSNHTQATN"  
 exon 1..246  
 /gene="HLA-DPA1"  
 /number=2

ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CGGACCATGTGTCAACTTATGCC 23

Db 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||

RESULT 3  
 HS0103X2  
 LOCUS 246 bp DNA linear PRI 14-JUL-1995  
 DEFINITION H.sapiens MHC class II HLA-DPA1\*0103 gene (exon 2).  
 X82390  
 VERSION X82390.1 GI:565028  
 KEYWORDS MHC class II HLA DPA1.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 246)  
 Roquemuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,  
 Marsh,S.G., Bodmer,J.G. and Tiliandus,M.G.  
 Sequencing-based typing reveals new insight in HLA-DPA1  
 Polymorphism  
 Tissue Antigens 45 (1), 57-62 (1995)  
 JOURNAL Location/Qualifiers

FEATURES  
 source 1..246  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /haplotype="DPA1\*0103"  
 /cell\_line="LB and LG2"  
 1..246  
 /gene="MHC class II HLA-DPA1 gene"  
 1..246  
 /gene="MHC class II HLA-DPA1 gene"  
 /number=2

gene  
 exon

ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 1 CGGACCATGTGTCAACTTATGCC 23

RESULT 4  
 HS0201X2  
 LOCUS 246 bp DNA linear PRI 14-JUL-1995  
 DEFINITION H.sapiens MHC class II HLA-DPA1\*0201 gene (exon 2).  
 X82394  
 VERSION X82394.1 GI:572739  
 KEYWORDS MHC class II HLA DPA1.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 246)  
 Roquemuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,  
 Marsh,S.G., Bodmer,J.G. and Tiliandus,M.G.  
 Sequencing-based typing reveals new insight in HLA-DPA1  
 Polymorphism  
 Tissue Antigens 45 (1), 57-62 (1995)  
 JOURNAL Location/Qualifiers

FEATURES  
 source 1..246  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /haplotype="DPA1\*0103"  
 /cell\_line="LB and LG2"  
 1..246  
 /gene="MHC class II HLA-DPA1 gene"  
 1..246  
 /gene="MHC class II HLA-DPA1 gene"  
 /number=2

```

7725312
REFERENCE 2 (bases 1 to 246)
AUTHORS Rozenmuller,E.H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) E.H. Rozenmuller, Diagnostic DNA Laboratory,
University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
Utrecht, NETHERLANDS
COMMENT Related sequences: S52453 and D14344.
Related sequences: S52453, D14344 and M83906.
FEATURES
    source
        1..246
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /haplotype="DPAI*0201"
            /cell_line="DAUDI and AKIBA"
            1..226
            /gene="MHC class II HLA-DPAI gene"
            1..226
            /gene="MHC class II HLA-DPAI gene"
            /number=2
    gene
    exon

ORIGIN
Query Match 100.0%; Score 23; DB 9; Length 246;
Best Local Similarity 100.0%; Pred.No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACTTATGCC 23
    |||||
Db 1 CGGACCATGTGTCACTTATGCC 23

RESULT 5
HSHLADPAI 246 bp DNA linear PRI 14-JUL-1995
LOCUS
DEFINITION H.sapiens HLA-DPAI gene, exon 2.
ACCESSION X78198
VERSION DPAI*01new; HLA-DPAI gene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Rozenmuller,E.H., Bouwens,A.G., van Oort,E., Versluis,L.F.,
Marsh,S.G., Bodmer,J.G. and Tilanus,M.G.:
Sequencing-based typing reveals new insight in HLA-DPAI
polymorphism
Tissue Antigens 45 (1), 57-62 (1995)
JOURNAL MEDLINE 95242313
PUBMED 7725312
REFERENCE 2 (bases 1 to 246)
AUTHORS Rozenmuller,E.H.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-1994) E.H. Rozenmuller, Diagnostic DNA Lab,
University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA.
Utrecht, NETHERLANDS
FEATURES
    source
        1..246
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /isolate="SK"
            /db_xref="taxon:9606"
            /chromosome="6"
            /haplotype="DPAI*01"
            1..246
            /gene="HLA-DPAI"
            1..246
            /gene="HLA-DPAI"
            /product="DPAI*01new"
            /number=2
    gene
    exon

ORIGIN

```

```

Query Match 100.0%; Score 23; DB 9; Length 246;
Best Local Similarity 100.0%; Pred.No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACTTATGCC 23
    |||||
Db 1 CGGACCATGTGTCACTTATGCC 23

RESULT 6
HSHZ48473 252 bp DNA linear PRI 06-APR-1998
LOCUS
DEFINITION H.sapiens HLA-DPAI gene for first domain of MHC class 2 molecule,
alpha-chain (allele DPAI*0203).
ACCESSION Z48473
VERSION Z48473.1 GI:1770743
KEYWORDS alpha-chain; MHC class 2 molecule.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muntau,B., Thyse,T., Pirmez,C. and Horstmann,R.D.
A novel DPAI allele (DPAI*0203) composed of known epitopes
Tissue Antigens 49 (6), 668-669 (1997)
9737898
MEDLINE 9234495
PUBMED
REFERENCE 2 (bases 1 to 252)
AUTHORS Muntau,B.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1995) Birgit Muntau, Molecular Genetics, Bernhard
Nocht Institute for Tropical Medicine, Hamburg,
Bernhard-Nocht-Str.74, 20359 Hamburg, Germany
On Jan 9, 1997 this sequence version replaced gi:683569.
COMMENT Location/Qualifiers
FEATURES
    source
        1..252
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /isolate="Brazilian White"
            /db_xref="taxon:9606"
            /clone="TC 48-3"
            /sex="female"
            /tissue_type="blood"
            /dev_stage="adult"
            31..252
            /gene="HLA-DPAI"
            31..252
            /gene="HLA-DPAI"
            /product="first domain of MHC class 2 molecule,
            alpha-chain"
            /note="allele DPAI*0203"
            /number=2
    gene
    exon

ORIGIN
Query Match 100.0%; Score 23; DB 9; Length 252;
Best Local Similarity 100.0%; Pred.No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACTTATGCC 23
    |||||
Db 8 CGGACCATGTGTCACTTATGCC 30

RESULT 7
AF165160 257 bp DNA linear PRI 01-SEP-2000
LOCUS
DEFINITION Homo sapiens MHC class II HLA-DPAI antigen (HLA-DPAI) gene,
HLA-DPAI*0201 variant allele, exon 2 and partial cds.
ACCESSION AF165160
VERSION AF165160.1 GI:5713147
KEYWORDS

```



SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 257)  
 AUTHORS McDaniel,D.O., Nguyen,C. and McDaniel,L.S.  
 TITLE A new HLA-DPA1 allele, DPA1\*02016, identified in African-American  
 population  
 JOURNAL Tissue Antigens 56 (2), 197-198 (2000)  
 MEDLINE 20470607  
 PUBMED 11019928  
 REFERENCE 2 (bases 1 to 257)  
 AUTHORS McDaniel,D.O., Nguyen,C. and McDaniel,L.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JUL-1999) Surgery, University of Mississippi Medical  
 Center, 2500 North State Street, Jackson, MS 39216, USA  
 FEATURES  
 source  
 1..257  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="peripheral blood cell"  
 <1..>257  
 /gene="HLA-DPA1"  
 /allele="HLA-DPA1\*0201 variant"  
 <1..>257  
 /gene="HLA-DPA1"  
 <1..>257  
 /gene="HLA-DPA1"  
 /codon\_start=1  
 /product="MHC class II HLA-DPA1 antigen"  
 /protein\_id="AAD47826.1"  
 /db\_xref="GI:5713148"  
 /translation="IKADHYSTYAAFYQTHRPTGEFMFEDEDEQFYVDLDRKETVWH  
 LEFGRAFSFAQGLANAILNNLTIQRSHNTQAANDP"  
 1..257  
 /gene="HLA-DPA1"  
 /number=2  
 ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 257;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 8 CGGACCATGTGTCAACTTATGCC 30  
 gene  
 mRNA  
 CDS  
 exon  
 ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 257;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 8 CGGACCATGTGTCAACTTATGCC 30  
 RESULT 8  
 AF118120  
 LOCUS AF118120 258 bp DNA linear PRI 24-MAR-1999  
 DEFINITION Homo sapiens MHC class II antigen DP alpha 1 subunit HLA-DPA1 gene  
 (bases 1 to 258)  
 ACCESSION AF118120  
 VERSION AF118120.1 GI:4469353  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 258)  
 AUTHORS McDaniel,D.O., Nguyen,C. and McDaniel,L.S.  
 TITLE A novel HLA-DPA1 variant DPA1\*02013 found in African-American  
 population  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 258)  
 AUTHORS McDaniel,D.O., Nguyen,C. and McDaniel,L.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JUL-1998) Surgery, UMMC, 2500 North State Street,  
 Clinical Science Bldg., Jackson, MS 39216, USA

COMMENT NCBI staff are still waiting for submitters to provide appropriate  
 coding region information.  
 FEATURES  
 source  
 1..258  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="peripheral blood mononuclear cells"  
 <1..>258  
 /gene="HLA-DPA1"  
 /note="MHC class II antigen DP alpha 1 subunit"  
 /allele="HLA-DPA1\*02013"  
 <1..>258  
 /gene="HLA-DPA1"  
 /number=2  
 ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 8 CGGACCATGTGTCAACTTATGCC 30  
 RESULT 9  
 AX237167  
 LOCUS AX237167 267 bp DNA linear PAT 26-SEP-2001  
 DEFINITION Sequence 143 from Patent WO0164886.  
 ACCESSION AX237167  
 VERSION AX237167.1 GI:15796721  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Gaiger,A., Algate,P.A. and Mannion,J.  
 TITLE Compositions and methods for the detection, diagnosis and therapy  
 of hematological malignancies  
 JOURNAL Patent: WO 0164886-A 143 07-SEP-2001;  
 CORIAX CORPORATION (US)  
 FEATURES  
 source  
 1..267  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 100.0%; Score 23; DB 6; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 129 CGGACCATGTGTCAACTTATGCC 151  
 RESULT 10  
 AX237352  
 LOCUS AX237352 267 bp DNA linear PAT 26-SEP-2001  
 DEFINITION Sequence 328 from Patent WO0164886.  
 ACCESSION AX237352  
 VERSION AX237352.1 GI:15796906  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1

AUTHORS Gaiger,A., Algate,P.A. and Mannion,J.  
 TITLE Compositions and methods for the detection, diagnosis and therapy of hematological malignancies  
 JOURNAL Patent: WO 0164886-A 328 07-SEP-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1..267  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 23; DB 6; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 129 CGGACCATGTGTCAACTTATGCC 151  
 |||||

RESULT 11  
 HSHLADPAX  
 LOCUS H.sapiens HLA-DPA1 gene. 268 bp DNA linear PRI 14-SEP-1995  
 DEFINITION H.sapiens HLA-DPA1 gene.  
 ACCESSION X83610  
 VERSION X83610.1 GI:987073  
 KEYWORDS HLA-DPA1 gene.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 Versluis,J.F., Verduyn,W., Van der Zwan,A., Oudshoorn,M. and Tilanus,M.G.J.  
 TITLE An update of the exon 2 sequence of the HLA-DPA1\*02012 allele  
 JOURNAL Tissue Antigens 46 (3 Pt 1), 206-207 (1995)  
 MEDLINE 96097411  
 PUBMED 8525481

REFERENCE 2 (bases 1 to 268)  
 Tilanus,M.G.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-DEC-1994) M.G.J. Tilanus, Diagnostic DNA Laboratory, Academic Hospital Utrecht, Heidelberglaan 100, PO Box 85500, 3508 GA Utrecht, NETHERLANDS  
 COMMENT Related Sequence: I31624.  
 FEATURES Location/Qualifiers  
 source 1..268  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /haplotype="DPA1\*02012"  
 <1..18  
 19..264  
 /gene="HLA-DPA1"  
 19..264  
 /gene="HLA-DPA1"  
 265..268

intron  
 gene  
 exon  
 intron

ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 19 CGGACCATGTGTCAACTTATGCC 41  
 |||||

RESULT 12  
 AX237066  
 LOCUS

AUTHORS Gaiger,A., Algate,P.A. and Mannion,J.  
 TITLE Compositions and methods for the detection, diagnosis and therapy of hematological malignancies  
 JOURNAL Patent: WO 0164886-A 328 07-SEP-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1..267  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 23; DB 6; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 134 CGGACCATGTGTCAACTTATGCC 156  
 |||||

RESULT 13  
 HSU87556  
 LOCUS Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2, partial cds.  
 DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2, partial cds.  
 ACCESSION U87556  
 VERSION U87556.1 GI:2760313  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 286)  
 Steiner,L., Begovich,A. and Suraj,V.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems, 1145 Atlantic Ave., Alameda, CA 94501, USA  
 COMMENT On Jan 8, 1998 this sequence version replaced gi:1842112.  
 FEATURES Location/Qualifiers  
 source 1..286  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 <1..286  
 /gene="HLA-D"  
 <16..261  
 /gene="HLA-D"  
 /note="MHC class II HLA-DPA1 antigen"  
 /codon\_start=3  
 /protein\_id="AAB97110.1"  
 /db\_xref="GI:2795772"  
 /translation="DHVSYAFAVQTRPTGFMFEFDEQFYVLDKKETVMHLEE  
 FGQAFSFEAQGGIAIAILNNLNTLIQRSNHTQATN"  
 16..261  
 /gene="HLA-D"  
 /number=2

exon

ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AX237066  
 LOCUS

DEFINITION Sequence 42 from Patent WO0164886.  
 ACCESSION AX237066  
 VERSION AX237066.1 GI:15796620  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 Gaiger,A., Algate,P.A. and Mannion,J.  
 TITLE Compositions and methods for the detection, diagnosis and therapy of hematological malignancies  
 JOURNAL Patent: WO 0164886-A 42 07-SEP-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1..272  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 23; DB 6; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 134 CGGACCATGTGTCAACTTATGCC 156  
 |||||

RESULT 13  
 HSU87556  
 LOCUS Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2, partial cds.  
 DEFINITION Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2, partial cds.  
 ACCESSION U87556  
 VERSION U87556.1 GI:2760313  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 286)  
 Steiner,L., Begovich,A. and Suraj,V.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems, 1145 Atlantic Ave., Alameda, CA 94501, USA  
 COMMENT On Jan 8, 1998 this sequence version replaced gi:1842112.  
 FEATURES Location/Qualifiers  
 source 1..286  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 <1..286  
 /gene="HLA-D"  
 <16..261  
 /gene="HLA-D"  
 /note="MHC class II HLA-DPA1 antigen"  
 /codon\_start=3  
 /protein\_id="AAB97110.1"  
 /db\_xref="GI:2795772"  
 /translation="DHVSYAFAVQTRPTGFMFEFDEQFYVLDKKETVMHLEE  
 FGQAFSFEAQGGIAIAILNNLNTLIQRSNHTQATN"  
 16..261  
 /gene="HLA-D"  
 /number=2

exon

ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 16 CGGACCATGTGTCAACTTATGCC 38

RESULT 14
AF015295 287 bp DNA linear PRI 29-SEP-1998
LOCUS Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*02013
DEFINITION allele), exon 2 and partial cds.
ACCESSION AF015295
VERSION AF015295.1 GI:3660653
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Steiner,L.L., Cavalli,A., Zimmerman,P.A., Boatn,B.A.,
TITANJ,V.P., Bradley,J.S., Lucius,R., Nutman,T.B. and
BEGOVICH,A.B.
Three new DP alleles identified in sub-Saharan Africa: DPB1*7401,
DPB1*02013, and DPA1*0302
JOURNAL Tissue Antigens 51 (6), 653-657 (1998)
MEDLINE 98357732
PubMed 3694359
REFERENCE 2 (bases 1 to 287)
AUTHORS Steiner,L., Begovich,A. and Zimmerman,P.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1997) Human Genetics, Roche Molecular Systems,
Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA
FEATURES
source
Location/Qualifiers
1..287
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
<1..>287
/gene="HLA-DPA1"
/allele="HLA-DPA1*02013"
<1..15
/gene="HLA-DPA1"
/number=1
<16..>261
/gene="HLA-DPA1"
/product="MHC class II antigen"
<16..>261
/gene="HLA-DPA1"
/codon_start=3
/product="MHC class II antigen"
/protein_id="AAC61669.1"
/db_xref="GI:3660654"
translation="DHVSTYAAFPVQTRPTGTFMFDEDEQFYVDLXKETVWHLLE
FGRAFSFEAQGGGLAILNNLTIQRSNHTQAA"
16..261
/gene="HLA-DPA1"
/number=2
262..>287
/gene="HLA-DPA1"
/number=2

exon
intron

Query Match 100.0%; Score 23; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 16 CGGACCATGTGTCAACTTATGCC 38

RESULT 15
AX237304/c

LOCUS AX237304 280 from Patent WO0164886.
DEFINITION Sequence 280 from Patent WO0164886.
ACCESSION AX237304
VERSION AX237304.1 GI:15796858
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Gaiger,A., Algate,P.A. and Mannion,J.
TITL Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A 280 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117

RESULT 16
AX237554/c
LOCUS AX237554 294 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 530 from Patent WO0164886.
ACCESSION AX237554
VERSION AX237554.1 GI:15797108
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Gaiger,A., Algate,P.A. and Mannion,J.
TITL Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A 530 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117

RESULT 17
AF013767
LOCUS AF013767 326 bp DNA linear PRI 15-OCT-1998
DEFINITION Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*0302
AUTHORS allele), partial cds.
ACCESSION AF013767
VERSION AF013767.1 GI:3660651
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

```

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 326)  
AUTHORS Steiner, L.L., Cavalli, A., Zimmerman, P.A., Boatín, B.A.,  
Titanji, V.P., Bradley, J.E., Lucius, R., Nutman, T.B. and  
Begovich, A.B.  
TITLE Three new DP alleles identified in sub-Saharan Africa: DPB1\*740-,  
DPAL\*02013, and DPAL\*0302  
JOURNAL Tissue Antigens 51 (6), 653-657 (1998)  
MEDLINE 98357732  
PubMed 9694359  
REFERENCE 2 (bases 1 to 326)  
AUTHORS Steiner, L., Begovich, A. and Zimmerman, P.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-1997) Human Genetics, Roche Molecular Systems,  
1145 Atlantic Ave., Alameda, CA 94501, USA  
COMMENT On Oct 16, 1998 this sequence version replaced gi:2865247.  
FEATURES  
source location/Qualifiers  
1..326  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
1..326  
/gene="HLA-DPA1"  
/allele="HLA-DPA1\*0302"  
1..326  
/gene="HLA-DPA1"  
/product="MHC class II antigen"  
1..326  
/gene="HLA-DPA1"  
/codon\_start=1  
/product="MHC class II antigen"  
/protein\_id="AAC64233.1"  
/db\_xref="GI:3757788"  
/translation="TFCKVFLYLYADHVSTYAMFVOTHPTEGFPEFDEMFYV  
DLDKKETVHLEEFQAFSFAQGLANAIINNNLTIQSNHTQATNGTPYLCLF  
LCSPTG"

Query Match 100.0%; Score 23; DB 9; Length 326;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
Db 38 CGGACCATGTGTCAACTTATGCC 60  
|||||

RESULT 18  
BD058253  
LOCUS secreted expressed sequence tags (sESTs).  
DEFINITION BD058253  
ACCESSION BD058253  
VERSION BD058253.1 GI:22603859  
KEYWORDS JP 2001519666-A/108.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 396)  
Tracy, M., Mccoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,  
Treacy, M., Spaulding, V. and Agostino, M.J.  
Secreted expressed sequence tags (sESTs)  
Patent: JP 2001519666-A 108 23-OCT-2001;  
GENETICS INSTITUTE INC  
PN JP 2001519666-A/108  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI

Query Match 100.0%; Score 23; DB 9; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
Db 38 CGGACCATGTGTCAACTTATGCC 60  
|||||

RESULT 18  
BD058253  
LOCUS secreted expressed sequence tags (sESTs).  
DEFINITION BD058253  
ACCESSION BD058253  
VERSION BD058253.1 GI:22603859  
KEYWORDS JP 2001519666-A/108.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 396)  
Tracy, M., Mccoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,  
Treacy, M., Spaulding, V. and Agostino, M.J.  
Secreted expressed sequence tags (sESTs)  
Patent: JP 2001519666-A 108 23-OCT-2001;  
GENETICS INSTITUTE INC  
PN JP 2001519666-A/108  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI

DAVID MERBERG.  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
1..396  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"

FEATURES  
source location/Qualifiers  
1..396  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"

ORIGIN  
Query Match 100.0%; Score 23; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
Db 152 CGGACCATGTGTCAACTTATGCC 174  
|||||

RESULT 19  
BD058339  
LOCUS secreted expressed sequence tags (sESTs).  
DEFINITION BD058339  
ACCESSION BD058339  
VERSION BD058339.1 GI:22603945  
KEYWORDS JP 2001519666-A/194.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 466)  
Jacobs, K., Mccoy, J.M., Lavallie, E.R., Racie, L.A., Verberg, D.,  
Treacy, M., Spaulding, V. and Agostino, M.J.  
Secreted expressed sequence tags (sESTs)  
Patent: JP 2001519666-A 194 23-OCT-2001;  
GENETICS INSTITUTE INC  
PN JP 2001519666-A/194  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI

DAVID MERBERG.  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
1..466  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"

FEATURES  
source location/Qualifiers  
1..466  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"

ORIGIN  
Query Match 100.0%; Score 23; DB 6; Length 466;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
Db 120 CGGACCATGTGTCAACTTATGCC 142  
|||||

RESULT 20  
AX884252  
LOCUS Sequence 115 from Patent EP1033401.  
DEFINITION AX884252  
ACCESSION AX884252  
VERSION AX884252.1 GI:40039227

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.  
AUTHORS Expressed sequence tags and encoded human proteins  
TITLE Patent: EP 1033401-A 115 06-SEP-2000;  
JOURNAL Genset (FR)  
FEATURES  
source Location/Qualifiers  
1..476  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
95..>475  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE98366.1"  
/db\_xref="GI:40039228"  
/translation="MRPEDRMFHRAVILRALSLAFLSLRGAGAIKADHVSTVAAPV  
QTHRPTEGFMEFDEDEMFYVLDKXETVWHLFEFGQAFSFEAGGLANIALNNLN  
TLIQRSNHTQATNDPPEVTFPKPEP"  
sig\_peptide 95..187  
/note="score 10.1 seq SLAFLSLRGAGA/IK"  
ORIGIN  
Query Match 100.0%; Score 23; DB 6; Length 476;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
DB 195 CGGACCATGTGTCAACTTATGCC 217  
|||||  
RESULT 21  
BD023862 476 bp DNA linear PAT 27-AUG-2002  
LOCUS Sequence tag and encoded human protein.  
DEFINITION BD023862  
ACCESSION BD023862  
VERSION BD023862.1 GI:22565085  
KEYWORDS JP 2001269182-A/108.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 476)  
REFERENCE Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
AUTHORS Sequence tag and encoded human protein  
TITLE Patent: JP 2001269182-A 108 02-OCT-2001;  
JOURNAL GENSET  
COMMENT OS Homo sapiens (human)  
PN JP 2001269182-A/108  
PD 02-OCT-2001  
PF 24-FEB-2000 JP 2000118773  
PR 26-FEB-1999 US 60/122487  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES  
PI JORDAN  
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC  
C12N5/10,  
PC C12P21/02, C12P21/08, C12Q1/69//G06F17/30, C12N15/00, C12N5/00, PC  
G06F15/40  
CC score 10.1  
CC seq SLAFLSLRGAGA/IK  
FH Key Location/Qualifiers  
FT CDS 95..475  
FT sig\_peptide 95..187.  
FT Location/Qualifiers  
1..476  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 23; DB 6; Length 476;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
DB 195 CGGACCATGTGTCAACTTATGCC 217  
|||||  
RESULT 22  
AX884251 576 bp DNA linear PAT 18-DEC-2003  
LOCUS Sequence 114 from Patent EP1033401.  
DEFINITION AX884251  
ACCESSION AX884251  
VERSION AX884251.1 GI:40039225  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.  
AUTHORS Expressed sequence tags and encoded human proteins  
TITLE Patent: EP 1033401-A 114 06-SEP-2000;  
JOURNAL Genset (FR)  
FEATURES  
source Location/Qualifiers  
1..576  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
195..>575  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE98365.1"  
/db\_xref="GI:40039226"  
/translation="MRPEDRMFHRAVILRALSLAFLSLRGAGAIKADHVSTVAAPV  
QTHRPTEGFMEFDEDEMFYVLDKXETVWHLFEFGQAFSFEAGGLANIALNNLN  
TLIQRSNHTQATNDPPEVTFPKPEP"  
sig\_peptide 195..287  
/note="score 10.1 seq SLAFLSLRGAGA/IK"  
ORIGIN  
Query Match 100.0%; Score 23; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
DB 295 CGGACCATGTGTCAACTTATGCC 317  
|||||  
RESULT 23  
BD023861 576 bp DNA linear PAT 27-AUG-2002  
LOCUS Sequence tag and encoded human protein.  
DEFINITION BD023861  
ACCESSION BD023861  
VERSION BD023861.1 GI:22565084  
KEYWORDS JP 2001269182-A/107.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 576)  
REFERENCE Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
AUTHORS Sequence tag and encoded human protein  
TITLE Patent: JP 2001269182-A 107 02-OCT-2001;  
JOURNAL GENSET  
COMMENT OS Homo sapiens (human)  
PN JP 2001269182-A/107  
PD 02-OCT-2001  
PF 24-FEB-2000 JP 2000118773

```

PR 26-FEB-1999 US 60/122487
PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN VVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC score 10.1
CC seq SLAFLLSLRGAGA/IK
CC key Location/Qualifiers
FT CDS 195..575
FT sig_peptide 195..287.
FT Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 23; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 295 CGGACCATGTGTCAACTTATGCC 317

RESULT 24
LOCUS E00485 661 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of a fragment of pSBalpha-318.
ACCESSION E00485
VERSION E00485.1 GI:2168768
KEYWORDS JP 1985226888-A/3.
SOURCE unidentified
ORGANISM unclassified.
1 (bases 1 to 661)
REFERENCE Edowado, R.S.Z., Kerii, B.M., Koorei, H.R. and Henrii, R.
AUTHORS NUCLEIC ACID LABELLING SUBSTANCE AND USE
TITLE Patent: JP 1985226888-A 3 12-NOV-1985;
JOURNAL CBTUS CORP
COMMENT OS Human {Homo sapiens}
PN JP 1985226888-A/3
PD 12-NOV-1985
PF 20-MAR-1985 JP 1985054705
PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 633263 PI
EDOWADO RUISU SHIERDON ZA SAADO, KERII BAKUSU MIYURISU, PI
KOOREI HAWAADO REBENSON, HENRII RAPOPORUTO
PC C07D519/00, C07H21/04, C12N15/00, C12Q1/68, G01N33/50, G01N33/532,
PC G01N33/58,
PC C07D519/00, C07D493:04, C07D495:04), (C07D519/00, C07D493:04, PC
C07D493:10);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue type=Blood;
CC *source: cell type=B cell;
CC *source: library=cDNA library;
FH Key Location/Qualifiers
FH misc_feature 1..661
FT /note='a fragment derived from pSBalpha-318
FT for insertion'.
FT Location/Qualifiers
1..661
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 23; DB 9; Length 690;
Best Local Similarity 100.0%; Pred. No. 0.13;

```

```

Query Match 100.0%; Score 23; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 179 CGGACCATGTGTCAACTTATGCC 201

RESULT 25
LOCUS I03086 661 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 5 from Patent US 4582789.
ACCESSION I03086
VERSION I03086.1 GI:268242
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
1 (bases 1 to 661)
REFERENCE Sheldon, E.L. III, Levenson, C.F., Mullis, K.B. and Rapoport, H.
AUTHORS Process for labeling nucleic acids using psoralen derivatives
TITLE Patent: US 4582789-A 5 15-APR-1986;
JOURNAL Cetus Corporation; Emeryville, CA
FEATURES Location/Qualifiers
source 1..661
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 23; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 179 CGGACCATGTGTCAACTTATGCC 201

RESULT 26
LOCUS S40633 690 bp DNA linear PRI 06-MAY-1993
DEFINITION HLA class II: DPAl (DPAl*0101) [human, Genomic, 690 nt].
ACCESSION S40633
VERSION S40633.1 GI:1679890
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 690)
REFERENCE Marsh, S.G. and Bodmer, J.G.
AUTHORS HLA class II nucleotide sequences, 1991
TITLE Immunogenetics 33 (5-6), 321-334 (1991)
JOURNAL 91267561
MEDLINE 1904836
PUBMED
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibseq 40633] from the original journal article.
This sequence comes from Figure 15.
On Nov 21, 1996 this sequence version replaced gi:1619630.
Region: HLA class II.
FEATURES Location/Qualifiers
source 1..690
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..690
/gene="DPAl"
/allele="DPAl*0101"

ORIGIN
Query Match 100.0%; Score 23; DB 9; Length 690;
Best Local Similarity 100.0%; Pred. No. 0.13;

```

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 8 CGGACCATGTGTCAACTTATGCC 30

RESULT 27  
 HUMHDPFA 818 bp mRNA linear PRI 07-MAR-1995  
 LOCUS Homo sapiens MEC class II DPw3-alpha-1 chain mRNA, complete cds.  
 DEFINITION M27487  
 ACCESSION M27487.1 GI:703088  
 VERSION  
 KEYWORDS cell surface glycoprotein; class II gene; integral membrane protein; lymphocyte antigen; major histocompatibility complex.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 818)  
 REFERENCE Young, J.A., Lindsay, J., Bodmer, J.G. and Trowsdale, J.  
 AUTHORS Epitope recognition by a DP alpha chain-specific monoclonal antibody (Dp11.1) is influenced by the interaction between the DP alpha chain and its polymorphic DP beta chain partner  
 TITLE Hum. Immunol. 23 (1), 37-44 (1988)

JOURNAL Hum. Immunol. 23 (1), 37-44 (1988)  
 MEDLINE 89053719  
 PUBMED 2461352

COMMENT On Mar 9, 1995 this sequence version replaced gi:341719.  
 Original location text: Homo sapiens cDNA to mRNA.  
 Location/Qualifiers  
 FEATURES  
 source 1..818  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="6921.3"  
 /cell\_line="LLC1CRF"  
 /cell\_type="B lymphoblast"

gene 1..818  
 /gene="HLA-DPA1"

CDS 36..818  
 /codon\_start=1  
 /gene="HLA-DPA1"  
 /product="MHC class II DP3-alpha"  
 /protein\_id="AAA63220.1"  
 /db\_xref="GI:703089"  
 /db\_xref="GDB:600-120-634"  
 /translation="MRPDRPHRAVILRALSLAFLSLRGAGAIKADHYSTVAAFV QTHRTGFMFEFDEDEMYDLDKKTWHLEFGQAFSFEAQGLANILNLLN LGQNTLIQNDPVTVPKPELVQPNLTICHIDKPPVLANVTWLCNGELVT EGVASLFLPRTDYSFKFHYLTFTVSAEDFYDCSVEHWGLDQPLKHWAEPIQMP ETTEIVLCALGLVLGLIIVGTVLIIKSLRSGHDPRAQGTL"

sig\_peptide 36..128  
 /gene="HLA-DPA1"  
 /note="G00-120-634"

mat\_peptide 129..815  
 /gene="HLA-DPA1"  
 /product="MHC class II DP3-alpha"  
 /note="G00-120-634"

ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 818;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 136 CGGACCATGTGTCAACTTATGCC 158

RESULT 28  
 HSSBAL 1048 bp mRNA linear PRI 13-DEC-1996  
 LOCUS Human mRNA for SB classII histocompatibility antigen alpha-chain.  
 DEFINITION

ACCESSION X00457.1 GI:36405  
 VERSION X00457.1  
 KEYWORDS antigen; membrane protein; signal peptide.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1048)  
 REFERENCE Strominger, J.L.  
 AUTHORS Isotypic and allotypic variation of human class II histocompatibility antigen alpha-chain genes  
 TITLE Nature 308 (5957), 327-333 (1984)  
 JOURNAL Nature 308 (5957), 327-333 (1984)  
 MEDLINE 84168117  
 PUBMED 6584734

COMMENT On Nov 6, 2003 this sequence version replaced gi:188516.  
 Location/Qualifiers  
 FEATURES  
 source 1..1048  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 <1..703  
 /codon\_start=2  
 /product="SB classII histocompatibility antigen alpha-chain"  
 /protein\_id="CAA25143.1"  
 /db\_xref="GI:758100"  
 /db\_xref="GOA:P20036"  
 /translation="GAGAIKADHYSTVAAFVQTHRTGFMFEFDEDEMYDLDKKE TVWHLERFGQAFSFEAQGLANILNLLNLTIIQNSHTQATNDPPEVTVPKPEVPS LGQNTLIQNDPVTVPKPELVQPNLTICHIDKPPVLANVTWLCNGELVTGVASLFLPRTDYSFKFHYLTFTVPS AEDFYDCRVEHWGLDQPLKHWAEPIQMPETTEIVLCALGLVLGLVGVGVGVLLII KSLRSGHDPRAQGTL"

sig\_peptide <1..113  
 /note="signal peptide fragment"

misc\_feature 14..265  
 /note="alpha 1"

misc\_feature 266..547  
 /note="alpha 2"

misc\_feature 548..586  
 /note="(CP) connecting peptide"

misc\_feature 587..655  
 /note="(TM) transmembrane region"

misc\_feature 656..700  
 /note="(Cy) cytoplasmatic region"

polyA\_site 1048  
 /note="polyadenylation site"

ORIGIN  
 Query Match 100.0%; Score 23; DB 9; Length 1048;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 Db 21 CGGACCATGTGTCAACTTATGCC 43

RESULT 29  
 I03088 1140 bp ss-DNA linear PAT 21-MAY-1993  
 LOCUS Sequence 7 from Patent US 4582789.  
 DEFINITION I03088  
 ACCESSION I03088  
 VERSION I03088.1 GI:268244  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 1 (bases 1 to 1140)  
 REFERENCE Shelton, E.L. III, Levenson, C.H., Mullis, K.B. and Rapoport, H.  
 AUTHORS Process for labeling nucleic acids using psoralen derivatives  
 TITLE Patent: US 4582789-A 7 15-APR-1986;  
 JOURNAL

```

Cetus Corporation; Emeryville, CA
FEATURES
  Source
    1..1140
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1140;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202
RESULT 30
E00484
LOCUS      E00484      1201 bp      RNA      linear      PAT 29-SEP-1997
DEFINITION DNA sequence of pSbalpha-318.
ACCESSION  E00484
VERSION     E00484.1 GI:2168767
KEYWORDS   JP 1985226888-A/2.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Edowaado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE     NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL    Patent: JP 1985226888-A 2 12-NOV-1985;
          CETUS CORP
COMMENT    OS Human (Homo sapiens)
          PN JP 1985226888-A/2
          PD 12-NOV-1985
          PF 20-MAR-1985 JP 1985054705
          PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI
          EDOWAAD RUISU SHERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
          KOCRAI HAWAADO REBENSON, HENRII RAPOPORUTO
          PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/50,G01N33/532,
          G01N33/58,
          PC (C07D519/00,C07D493:04,C07D495:04),(C07D519/00,C07D493:04, PC
          C07D493:10);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          CC *source: tissue type=Blood;
          CC *source: cell type=B cell;
          CC *source: library=cDNA library;
          CC *source: clone=pSbalpha-318;
          FH Key      Location/Qualifiers
          FT CDS      1..2201
          FT          /gene='pSbalpha-318'.
          FT          Location/Qualifiers
          FT          1..1201
          /organism="unidentified"
          /mol_type="genomic RNA"
          /db_xref="taxon:32644"
FEATURES
  Source
    1..1201
    /organism="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1201;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202
RESULT 31
E00484
LOCUS      E00484      1201 bp      ss-DNA      linear      PAT 21-MAY-1993
DEFINITION DNA sequence of pSbalpha-318.
ACCESSION  E00484
VERSION     E00484.1 GI:2168767
KEYWORDS   JP 1985226888-A/2.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Edowaado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE     NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL    Patent: JP 1985226888-A 2 12-NOV-1985;
          CETUS CORP
COMMENT    OS Human (Homo sapiens)
          PN JP 1985226888-A/2
          PD 12-NOV-1985
          PF 20-MAR-1985 JP 1985054705
          PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI
          EDOWAAD RUISU SHERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
          KOCRAI HAWAADO REBENSON, HENRII RAPOPORUTO
          PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/50,G01N33/532,
          G01N33/58,
          PC (C07D519/00,C07D493:04,C07D495:04),(C07D519/00,C07D493:04, PC
          C07D493:10);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          CC *source: tissue type=Blood;
          CC *source: cell type=B cell;
          CC *source: library=cDNA library;
          CC *source: clone=pSbalpha-318;
          FH Key      Location/Qualifiers
          FT CDS      1..2201
          FT          /gene='pSbalpha-318'.
          FT          Location/Qualifiers
          FT          1..1201
          /organism="unidentified"
          /mol_type="genomic RNA"
          /db_xref="taxon:32644"
FEATURES
  Source
    1..1201
    /organism="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1201;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202
RESULT 31
E00484
LOCUS      E00484      1201 bp      ss-DNA      linear      PAT 21-MAY-1993
DEFINITION DNA sequence of pSbalpha-318.
ACCESSION  E00484
VERSION     E00484.1 GI:2168767
KEYWORDS   JP 1985226888-A/2.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Edowaado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE     NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL    Patent: JP 1985226888-A 2 12-NOV-1985;
          CETUS CORP
COMMENT    OS Human (Homo sapiens)
          PN JP 1985226888-A/2
          PD 12-NOV-1985
          PF 20-MAR-1985 JP 1985054705
          PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI
          EDOWAAD RUISU SHERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
          KOCRAI HAWAADO REBENSON, HENRII RAPOPORUTO
          PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/50,G01N33/532,
          G01N33/58,
          PC (C07D519/00,C07D493:04,C07D495:04),(C07D519/00,C07D493:04, PC
          C07D493:10);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          CC *source: tissue type=Blood;
          CC *source: cell type=B cell;
          CC *source: library=cDNA library;
          CC *source: clone=pSbalpha-318;
          FH Key      Location/Qualifiers
          FT CDS      1..2201
          FT          /gene='pSbalpha-318'.
          FT          Location/Qualifiers
          FT          1..1201
          /organism="unidentified"
          /mol_type="genomic RNA"
          /db_xref="taxon:32644"
FEATURES
  Source
    1..1201
    /organism="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1201;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202

```

```

Sequence 6 from Patent US 4617261.
ACCESSION  I03006
VERSION     I03006.1 GI:268462
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and
          Watson,R.M.
TITLE     Process for labeling nucleic acids and hybridization probes
JOURNAL    Patent: US 4617261-A 6 14-OCT-1986;
          Cetus Corporation; Emeryville, CA
FEATURES
  Source
    1..1201
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1201;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202
RESULT 32
I03423
LOCUS      I03423      1201 bp      ss-DNA      linear      PAT 21-MAY-1993
DEFINITION Sequence 6 from Patent US 4822731.
ACCESSION  I03423
VERSION     I03423.1 GI:270023
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Watson,R.M., Sheldon,E.L. III and Shead,R.M.
TITLE     Process for labeling single-stranded nucleic acids and
          hybridization probes
JOURNAL    Patent: US 4822731-A 6 18-APR-1989;
          Cetus Corporation; Emeryville, CA
FEATURES
  Source
    1..1201
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
  Query Match      100.0%; Score 23; DB 6; Length 1201;
  Best Local Similarity 100.0%; Pred. No. 0.14;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 180 CGGACCATGTGTCAACTTATGCC 202
RESULT 33
BC009956
LOCUS      BC009956      1201 bp      mRNA      linear      PRI 11-DEC-2003
DEFINITION Homo sapiens major histocompatibility complex, class II, DP alpha
          1, mRNA (cDNA clone MGC:14114 IMAGE:4309471), complete cds.
ACCESSION  BC009956
VERSION     BC009956.2 GI:39644570
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

```



Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuller,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,E., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marisina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., McQuellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Sosak,S.A., McEwan,P.J., McKerran,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Pahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

12477932

2 (bases 1 to 1201)

Strausberg,R.

Direct Submission

Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncbi.nih.gov>

On Dec 9, 2003 this sequence version replaced gi:14602922.

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Louis Staudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nsc.nih.gov/>

Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.B., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W., Tsurgonev,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>

Series: IRAL Plate: 20 Row: p Column: 4

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24797073.

Location/Qualifiers

#### FEATURES

1..1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:14114 IMAGE:4309471"

/tissue\_type="Primary B-Cells from Tonsils"

/clone\_lib="NIH MGC-48"

/lab\_host="DHIOE-R"

/note="Vector: pOTB7"

1..1201

/gene="HLA-DPA1"

/note="synonyms: HLA-DP1A, HLA-DP, HLA-SB"

/db\_xref="LocusID:3113"

/db\_xref="MIM:142880"

32..814

/codon\_start=1

/product="HLA-DPA1 protein"

/protein\_id="AAH09556.1"

#### gene

#### CDS

/db\_xref="GI:14602923"

/db\_xref="LocusID:3113"

/translation="MRPEDRMFHRAVILRALSLAFLLSLRGAGAKADIVSTYAFV QTRPTGEFMFEDEQFYVDLKKETVHLEEFQAFSFEAGGLANAILNNLN TLTRQSNTOAANDPPEVTFKPEVLPQNTLICHIDFFPVLNVLKNGEPVT EGVASLFLPRTDYSFKPHYLTFFVSAEDVYDCRVHGLDQPLLKHEAQBFQMP ETTETVLGALGLVGLVIGVIVLIIKLSRSHDPAQGL"

134..376

/note="MHC II alpha; Region: Class II histocompatibility antigen, alpha domain"

/db\_xref="CDD:pfam00993"

374..661

/note="IGC; Region: Immunoglobulin domain constant region subfamily"

/db\_xref="CDD:cd00098"

ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 1201;

Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23

Db 132 CGGACCATGTGTCAACTTATGCC 154

RESULT 34

AX552229

LOCUS AX552229 1259 bp DNA linear PAT 27-NOV-2002

DEFINITION Sequence 138 from Patent WO0162927.

ACCESSION AX552229

VERSION AX552229.1 GI:25896467

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1

AUTHORS Barville,S.C., Greenwalt,L.B., Lincoln,S.E., Stockdreher,T.K., Anshey,S., Chang,S.C., Chen,W., D'Sa,S.A., Dam,T.C., Liu,T.F., Rosen,B.H., Russo,F.D., Spiro,P.A., Bradley,D.L., Chen,A., Cohen,H.J., Daffo,A., Daniels,S.E., Dufour,G.B., Flores,V., Fong,W.T., Hodgson,D.M., Jackson,S., Jones,A.L., Panzer,S., Roseberry,A.M., Shah,P., Wright,R.J., Yap,P.E., Yu,J.Y., Bratcher,S.R., Chatur,M.S., Bahl,C.R. and Hillman,J.L.

TITLE Polypeptides and corresponding polynucleotides for diagnostics and therapeutics

JOURNAL Patent: WO 0162927-A 138 30-AUG-2001; Incyte Genomics, Inc. (US)

FEATURES

source Location/Qualifiers

1..1259

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

/note="Incyte ID No: LI:1169865.1:2000MAY01"

ORIGIN

Query Match 100.0%; Score 23; DB 6; Length 1259;

Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23

Db 315 CGGACCATGTGTCAACTTATGCC 337

RESULT 35

AX780118

LOCUS AX780118 1480 bp DNA linear PAT 14-JUL-2003

DEFINITION Sequence 2275 from Patent WO03039443..

ACCESSION AX780118

VERSION AX780118.1 GI:32697112

KEYWORDS

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
TITLE       Novel genetic markers for leukemias
JOURNAL     Patent: WO 03039443-A 2275 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE);
            Ludwig-Maximilian-Universitaet Muenchen (DE);
            PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES    Location/Qualifiers
            1..1480
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 23; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACCTTATGCC 23
|||||
Db 389 CGGACCATGTGTCACCTTATGCC 411

RESULT 36
LOCUS      AX780119          1480 bp      DNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 2276 from Patent WO03039443.
ACCESSION  AX780119
VERSION     AX780119.1 GI:32697113
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
TITLE       Novel genetic markers for leukemias
JOURNAL     Patent: WO 03039443-A 2276 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE);
            Ludwig-Maximilian-Universitaet Muenchen (DE);
            PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES    Location/Qualifiers
            1..1480
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 23; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACCTTATGCC 23
|||||
Db 389 CGGACCATGTGTCACCTTATGCC 411

RESULT 37
LOCUS      HUMHDC02          2986 bp      DNA      linear      PRI 07-JAN-1995
DEFINITION Human MHC class II lymphocyte antigen (DPw4-alpha-1) gene, exons
AUTHORS     M23904 J02738 M15446
M23904.1 GI:188385
KEYWORDS    cell surface glycoprotein; class II gene; integral membrane
            protein; major histocompatibility complex.
SEGMENT     2 of 3

```

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Gustafsson, K., Widmark, E., Jonsson, A.K., Serenius, B., Sachs, D.H.,
Larhammar, D., Rask, L. and Peterson, P.A.
TITLE       Class II genes of the human major histocompatibility complex.
            Evolution of the DP region as deduced from nucleotide sequences of
            the four genes
JOURNAL     J. Biol. Chem. 262 (18), 8778-8786 (1987)
MEDLINE     87250502
PUBMED      3036829
COMMENT     source text: Human T-cell DNA, clone p412-1 and clone
            p2703-1.
FEATURES    Location/Qualifiers
            1..2986
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            join(M23903.1:563..1095,1..2986)
            /gene="DPw4-alpha-1"
            order(M23903.1:663..1095,1..769)
            /gene="DPw4-alpha-1"
            /number=1
            770..1015
            /gene="DPw4-alpha-1"
            /number=2
            1016..1355
            /gene="DPw4-alpha-1"
            /number=2
            1356..1637
            /gene="DPw4-alpha-1"
            /number=3
            1638..1851
            /gene="DPw4-alpha-1"
            /number=3
            1852..2018
            /gene="DPw4-alpha-1"
            /number=4
            2019..2986
            /gene="DPw4-alpha-1"
            /note="does not fit consensus"
            /number=4
ORIGIN      About 2.4 kb after segment 1; chromosome 6p21.3.
Query Match      100.0%; Score 23; DB 9; Length 2986;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACCTTATGCC 23
|||||
Db 770 CGGACCATGTGTCACCTTATGCC 792

RESULT 38
LOCUS      HSHLASBA          14646 bp      DNA      linear      PRI 16-FEB-1995
DEFINITION Human HLA-SB (DP) alpha gene.
ACCESSION    X03100
VERSION      X03100.1 GI:32243
KEYWORDS     antigen; cell surface glycoprotein; class II antigen; glycoprotein;
            inverted repeat; Kpn repetitive sequence; major histocompatibility
            complex; repetitive sequence.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Lawrance, S.K., Das, H.K., Pan, J. and Weissman, S.M.
TITLE       The genomic organisation and nucleotide sequence of the HLA-SB (DP)
            alpha gene
JOURNAL     Nucleic Acids Res. 13 (20), 7515-7528 (1985)

```

```

MEDLINE      86041930
PUBMED       2997750
COMMENT      Data kindly reviewed (05-MAY-1987) by S.K. Lawrence.
FEATURES
  source      1. .14646
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /complement(1. .247)
              /gene="HLA-SB beta"
              /complement(<1. .247)
              /gene="HLA-SB beta"
              /complement(1. .78)
              /gene="HLA-SB beta"
              /number=1
              /complement(79. .247)
              /gene="HLA-SB beta"
              /number=1
              /complement(<79. .178)
              /gene="HLA-SB beta"
              /codon_start=1
              /protein_id="CAA26886.1"
              /db_xref="GI:32244"
              /translation="MWVLQVSAAPRTVALTALIMVLLTSVVQGRATP"
              /complement(323. .374)
              /note="beta consensus sequence, put. regulatory region"
              514. .526
              /note="imp. inverted repeat a"
              2189. .2200
              /note="imp. inverted repeat a'"
              2461. .2512
              /note="alpha consensus sequence; put. regulatory region"
              2571. .11200
              /gene="HLA-SB"
              join(2571. .2749,6334. .6579,6920. .7201,7416. .7582,
              10872. .11200)
              /gene="HLA-SB alpha"
              join(2571. .2749,6334. .6579,6920. .7201,7416. .7582,
              10872. .11200)
              /gene="HLA-SB alpha"
              2571. .11200
              /gene="HLA-SB"
              2571. .2749
              /gene="HLA-SB"
              /number=1
              join(2650. .2749,6334. .6579,6920. .7201,7416. .7570)
              /gene="HLA-SB alpha"
              /codon_start=1
              /product="class II antigen"
              /protein_id="CAA26887.1"
              /db_xref="GI:673417"
              /db_xref="GOA:P20036"
              /db_xref="SWISS-PROT:P20036"
              /translation="MRPEDRMFHRAVILRLSLAFLSLRGAGAIKADHVSTYAAFV
              QTHRPTGRMFREDEDESVYDLKKTWHLERFQAESEPAQGLAIALLNNVLN
              TLIORSNTQATNDPPKTVPKPEYELGQPNLICHDKFPFPPVNTVLCNGBLVT
              BGVAESLPRIDYSFHFPHLYTFVPSAEFDYDCRHHWGLDQPLLKHWEAQEPIQMP
              EETVFLCALGLGVGIIVGTVLIILKSLRSGHPRAQGTI"
              2650. .2742
              /gene="HLA-SB"
              2650. .2749
              /gene="HLA-SB"
              /note="precursor fragment"
              2724. .2725
              /gene="HLA-SB"
              /note="pot. alternate signal sequence splice site"
              2750. .5333
              /gene="HLA-SB"
              /number=1
              4964. .4990
              /gene="HLA-SB"
              /note="inverted repeat b"
repeat_unit  5063. .5088
              /gene="HLA-SB"
              /note="inverted repeat b'"
6334. .6579
              /gene="HLA-SB"
              /number=2
6580. .6919
              /gene="HLA-SB"
              /number=2
6920. .7201
              /gene="HLA-SB"
              /number=3
7202. .7415
              /gene="HLA-SB"
              /number=3
7416. .7582
              /gene="HLA-SB"
              /number=4
7583. .10871
              /gene="HLA-SB"
              /number=4
8601. .9100
              /gene="HLA-SB"
              /note="sequence homologous to IgC epsilon genes"
8991. .9029
              /gene="HLA-SB"
              /note="inverted repeat C"
10516. .10554
              /gene="HLA-SB"
              /note="inverted repeat C'"
10872. .11200
              /gene="HLA-SB"
              /number=5
              /note="Kpn repetitive sequence"
12301. .12800
ORIGIN
Query Match      100.0%; Score 23; DB 9; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CGGACCATGTGTCAACTTATGCC 23
          |||||
Db      6334 CGGACCATGTGTCAACTTATGCC 6356
RESULT 39
BX120009/c
LOCUS      BX120009
DEFINITION Human DNA sequence from clone DASS-227B13 on chromosome 6, complete
            sequence.
ACCESSION  BX120009
VERSION    BX120009.10 GI:33504469
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 64380)
            Sycamore,N.
            Direct Submission
            Submitted (07-AUG-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Aug 7, 2003 this sequence version replaced gi:31335527.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -----
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations

```

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC> Xxbac-227B13 is from a DNA-arts S8TO human bac library VECTOR: pBelOBAC11.

## FEATURES

source

Location/Qualifiers  
1..64380  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DASS-227B13"  
/clone\_lib="DNA-arts-BAC.1-S8TO.1"

## ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 64380;  
Best Local Similarity 100.0%; Pred. No. 0.21; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACACTTATGCC 23

Db 2137 CGGACCATGTGTCACACTTATGCC 2115

RESULT 40

AL8059:3/c

LOCUS

DEFINITION Human DNA sequence from clone Xxbac-22D21 on chromosome 6, complete sequence. PRI 24-OCT-2002

ACCESSION

AL805913

VERSION

AL805913.4

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Tracey A.

Direct Submission

Submitted (23-Oct-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)

On Oct 25, 2002 this sequence version replaced gi:2204654.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC> Xxbac-22D21 is from a DNA-arts QBL human bac library VECTOR: pBelOBAC11.

## FEATURES

source

Location/Qualifiers  
1..106728  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="XXbac-22D21"  
/clone\_lib="DNA-arts-BAC.1-QBL.1"

## ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 106728;  
Best Local Similarity 100.0%; Pred. No. 0.22; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACACTTATGCC 23

Db 95141 CGGACCATGTGTCACACTTATGCC 95119

RESULT 41

AL645931/c

LOCUS

DEFINITION Human DNA sequence from clone Xxbac-138A21 on chromosome 6, complete sequence. PRI 24-APR-2002

ACCESSION

AL645931

VERSION

AL645931.7

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Almeida, J.

Direct Submission

Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)

On Mar 21, 2002 this sequence version replaced gi:19031691.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality &gt;=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em:, EMBL; Sw:,

SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep/XXbac-138A21](http://www.sanger.ac.uk/Projects/C_elegans/wormpep/XXbac-138A21) is from a CHORI-501 human bac - 2GP cell line library VECTOR:

PTABAC2.1  
This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators.  
Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6/MHC>.

Location/Qualifiers

FEATURES  
source

```
1..124899
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="6"
  /clone="XXbac-138A21"
  /clone_lib="CHORI-501"
```

ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 124899;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 77822 CGGACCATGTGTCAACTTATGCC 77800

RESULT 42

AC011086/c

LOCUS

DEFINITION Homo sapiens chromosome 6 clone RP11-93F3 map 6, WORKING DRAFT

AC011086

AC011086.5 GI:10047675

VERSION HTG: HTGS\_PHASE1; HTGS\_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 181228)

Siiren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 6, clone RP11-93F3

Unpublished

2 (bases 1 to 181228)

Siiren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhalter,B.,

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., Dearellano,K., Dewar,K., Domino,M., Doneilan,L., Doyle,M.,

Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,

Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,

McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 181228)

Siiren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Beda,P., Boguslavskiy,L.,

Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,

Cheepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,

Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,

FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,

Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,

Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,

Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,

Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,

McPheters,R., Meldrim,J., Meneus,L., Milova,T., Mlenga,V.,

Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,

O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sognes,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 10, 2000 this sequence version replaced gi:7717099.

All repeats were identified using RepeatMasker.

Smt, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIPR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information:

Center project name: 93\_F\_3

Center Clone name: 11322

----- Summary Statistics

Sequencing Vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 166135 bases at least Q40

Consensus quality: 173448 bases at least Q30

Consensus quality: 176936 bases at least Q20

Insert size: 174000; agarose-fp

Insert size: 180128; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases.

NOTE: This is a 'working draft' sequence. It currently  
consists of 12 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

\* 1 27639: contig of 27639 bp in length

\* 27640 27739: gap of 100 bp

\* 27740 28970: contig of 1231 bp in length

\* 28971 29070: gap of 100 bp

\* 29071 82294: contig of 53224 bp in length

\* 82295 82394: gap of 100 bp

\* 82395 87214: contig of 4820 bp in length

\* 87215 93029: contig of 5715 bp in length

\* 93030 101779: gap of 100 bp

\* 93130 101779: contig of 8650 bp in length

\* 101780 101879: gap of 100 bp

\* 101880 103114: contig of 6235 bp in length

\* 103115 108214: gap of 100 bp

\* 108215 116143: contig of 7929 bp in length

\* 116144 116244: gap of 100 bp

\* 116244 133838: contig of 17595 bp in length

\* 133839 133938: gap of 100 bp

\* 133939 153417: contig of 19479 bp in length

\* 153418 15517: gap of 100 bp

\* 15518 178444: contig of 24927 bp in length

\* 178445 178544: gap of 100 bp

\* 178545 181228: contig of 2684 bp in length.

\* Location/Qualifiers

1..181228

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="6"

/map="6"

FEATURES  
source

```

misc_feature      /clone="RP11_93F3"
                  /clone_lib="RPC1-11 Human Male BAC"
                  1..27639
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:left"
misc_feature      27740..28970
                  /note="assembly_fragment"
                  29071..82294
                  /note="assembly_fragment"
misc_feature      82395..87214
                  /note="assembly_fragment"
misc_feature      87315..93029
                  /note="assembly_fragment"
misc_feature      93130..101779
                  /note="assembly_fragment"
misc_feature      101880..108114
                  /note="assembly_fragment"
misc_feature      108215..116143
                  /note="assembly_fragment"
misc_feature      116244..133838
                  /note="assembly_fragment"
misc_feature      133939..153417
                  /note="assembly_fragment"
misc_feature      153518..178444
                  /note="assembly_fragment"
misc_feature      178545..181228
                  /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"

ORIGIN
Query Match      100.0%; Score 23; DB 2; Length 181228;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      : CGGACCATGTGTCAACTTATGCC 23
          |||||
Db      64060 CGGACCATGTGTCAACTTATGCC 64038

RESULT 43
AL662824/c
LOCUS           187964 bp      DNA      linear      PRI 24-APR-2002
DEFINITION      Human DNA sequence from clone Xxbac-22116 on chromosome 6, complete
sequence.
ACCESSION       AL662824
VERSION         AL662824.9 GI:20068657
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS         Almeida,J.
TITLE           Direct Submission
JOURNAL         Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
COMMENT         On Apr 7, 2002 this sequence version replaced gi:20067828.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following

```

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Xxbac-22116 is from a CHORI-502 human bac - COX cell line library VECTOR: PTARBAC2.1

This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC>.

## FEATURES

```

source
1..187964
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="6"
   /clone="Xxbac-22116"
   /clone_lib="CHORI-502"

```

## ORIGIN

```

Query Match      100.0%; Score 23; DB 9; Length 187964;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CGGACCATGTGTCAACTTATGCC 23
          |||||
Db      51851 CGGACCATGTGTCAACTTATGCC 51829

```

## RESULT 44

```

HUMXHCIIAX
LOCUS           214 bp      DNA      linear      PRI 07-JUN-1994
DEFINITION      Human MHC class II (HLA-DPA1) Gene.
ACCESSION       L11641
VERSION         L11641.1 GI:187872
KEYWORDS        alpha-subunit; class II gene; major histocompatibility complex.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS         Guechlein,L.A., Bias,W.B. and Schmeckpeper,B.J.
TITLE           New DP sequences: three DPA1 and one DPB1
JOURNAL         Tissue Antigens 41 (5), 269-272 (1993)
MEDLINE         94054425
PUBMED          8236240
COMMENT         Original source text: Homo sapiens DNA.
FEATURES
source
1..214
   /organism="Homo sapiens"
   /mol_type="unassigned DNA"
   /db_xref="taxon:9606"
   /cell_line="KT17 (10W5024)"

```

## ORIGIN

```

Query Match      95.7%; Score 22; DB 9; Length 214;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 CGGACCATGTGTCAACTTATGCC 23
          |||||
Db      1 CGGACCATGTGTCAACTTATGCC 22

```

## RESULT 45

```

AF098794
LOCUS           244 bp      DNA      linear      PRI 01-SEP-2000
DEFINITION      Homo sapiens MHC class II antigen (HLA-DPA1) gene, exon 2 and
partial cds.
ACCESSION       AF098794
VERSION         AF098794.1 GI:3859561
KEYWORDS

```







# AUTHORS

Williams, S.  
 Direct Submission  
 Submitted (19-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: hunquery@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

## COMMENT

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
 Group in collaboration with Jethro Herberg and John Trowsdale.  
 Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6>  
 IMPORTANT: This sequence is not  
 the entire insert of clone XX-CD0826Q1. It may be shorter because we  
 sequence overlapping sections only once, except for a short  
 overlap.

The true left end of clone XX-IC11 is at 22748 in this sequence.  
 The true right end of clone RF5-1033510 is at 100 in this sequence.

## FEATURES

### source

1..22847  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XX-CD0826Q1"

### repeat\_region

1..579  
 /note="L1PB1 repeat: matches 5524..6145 of consensus"

### repeat\_region

693..837  
 /note="L12 repeat: matches 1290..1435 of consensus"

### repeat\_region

1200..1510  
 /note="L1MA8 repeat: matches 5999..6288 of consensus"

### repeat\_region

1761..2284  
 /note="L12 repeat: matches 2120..2680 of consensus"

### misc\_feature

complement(2291..2721)

### gene

join(3549..3685,3801..3970,4459..4610,6424..6468,  
 6594..6653,6760..6797)

### CDS

/pseudo  
 join(3549..3685,3801..3970,4459..4610,6424..6468,  
 6594..6653,6760..6797)  
 /gene="CD0826Q1.1"

/note="CD0826Q1.1 (pseudogene similar to part of collagen

type XI alpha 2 (COL11A2))

match: proteins: Tr:Q59866 Sw:P13942 Tr:Q61432 Sw:Q64739

Sw:P12107 Tr:Q90589 Tr:Q88207 Sw:Q61245 Sw:P20908

Tr:Q62750 Tr:Q57580 Tr:Q921WQ Sw:Q62280"

### /pseudo

/codon\_start=1

/evidence=not\_experimental

4186..4351

/note="MIR repeat: matches 65..261 of consensus"

### repeat\_region

6951..7315

/note="L1PB2 repeat: matches 5776..6154 of consensus"

### misc\_feature

7257..7409  
 /note="match: GSS: Em:AQ868440"

### repeat\_region

7399..8718  
 /note="L1PB2 repeat: matches 4435..5791 of consensus"

### repeat\_region

8719..9679  
 /note="MER11C repeat: matches 1..1067 of consensus"

### repeat\_region

9686..12936  
 /note="L1PB2 repeat: matches 1099..4439 of consensus"

### repeat\_region

12935..13961  
 /note="L1PB2 repeat: matches 1207..194 of consensus"

### gene

join(13269..13323,17426..17782,18163..18444,18660..18810)

### /gene="HLA-DPA2"

/pseudo

join(13269..13323,17426..17782,18163..18444,18660..18810)

### /gene="HLA-DPA2"

II, DP alpha 2 (pseudogene)

match: proteins: Tr:Q30181 Sw:P20036 Sw:P20755 Tr:Q31287

Tr:Q30437 Tr:Q19810 Tr:Q19434 Tr:Q30761 Tr:Q31064

Tr:Q31065 Tr:Q30302 Sw:P01906 Tr:Q07455

match: genomic DNA: Em:X01634"

### /pseudo

/codon\_start=1

/evidence=not\_experimental

13962..14259

/note="AluX repeat: matches 2..298 of consensus"

14260..14500

/note="L1PB2 repeat: matches 1449..1207 of consensus"

14768..14894

/note="FLAMC repeat: matches 1..127 of consensus"

15580..15845

/note="L2 repeat: matches 1995..2283 of consensus"

17042..17103

/note="MIR repeat: matches 79..144 of consensus"

17340..17171

/note="L2 repeat: matches 2703..2734 of consensus"

17585..17806

/note="2 copies 111 mer 100% conserved"

18959..19182

/note="MIR repeat: matches 11..262 of consensus"

19187..19339

/note="L1P repeat: matches 4059..4223 of consensus"

19433..19937

/note="match: GSS: Em:AQ814971"

19568..19896

/note="match: STS: Em:Q55743

match: GSS: Em:AQ323676"

19597..19954

/note="L1M4 repeat: matches 3345..3710 of consensus"

complement(19680..20106)

/note="match: GSS: Em:AQ297034"

20063..20156

/note="47 copies 2 mer to 62% conserved"

20704..20754

/note="L1ME3 repeat: matches 6030..6080 of consensus"

20808..20873

/note="L2 repeat: matches 2675..2739 of consensus"

20933..21171

/note="M1L1 repeat: matches 72..323 of consensus"

## ORIGIN

Query Match 81.7%; Score 18.8; DB 9; Length 22847;

Best Local Similarity 90.9%; Pred.No. 47;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACTATGCG 22

|||||

DB 17426 CAGACCATGTGTCAACATATGC 17447

## RESULT 52

AL845446/c

LOCUS

DEFINITION

Human DNA sequence from clone XXbac-79p13 on chromosome 6, complete

sequence.

ACCESSION

AL845446

VERSION

AL845446.7 GI:24474482

**KEYWORDS**  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE**  
**AUTHORS** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**TITLE** Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**JOURNAL** 1 (bases 1 to 93842)  
**COMMENT** Johnson, C.  
 Direct Submission  
 Submitted (31-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Nov 1, 2002 this sequence version replaced gi:24430255.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 from a CHORI-501 human bac - PGF cell line library VECTOR:  
 PTAREAC2.1  
 This sequence was generated from part of bacterial clone contigs  
 constructed by the MHC Haplotype Consortium and collaborators.  
 Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6/MHC>.  
**FEATURES**  
 source  
 1..93842  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXbac-79p13"  
 /clone\_lib="DNA-arts-BAC.1-QBL.1"  
**ORIGIN**  
 Query Match 81.7%; Score 18.8; DB 9; Length 93842;  
 Best Local Similarity 90.9%; Pred. No. 54;  
 Matches 20; Conservative 0; Mismatches 2; Indels 3; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTATGC 22  
 DB 13666 CAGACCATGTGTCAACTATGC 13645  
**RESULT 53**  
**AL645940/C**  
**LOCUS** Human DNA sequence from clone XXbac-157A10 on chromosome 6,  
**DEFINITION** complete sequence.  
**ACCESSION** AL645940  
**VERSION** AL645940.4 GI:18564736  
**KEYWORDS** HTG.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**  
**AUTHORS** Leongamornlert, D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Feb 6, 2002 this sequence version replaced gi:18070925.  
**COMMENT** During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 from a CHORI-501 human bac - PGF cell line library VECTOR:  
 PTAREAC2.1  
 This sequence was generated from part of bacterial clone contigs  
 constructed by the MHC Haplotype Consortium and collaborators.  
 Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6/MHC>.  
**FEATURES**  
 source  
 1..158033  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXbac-157A10"  
 /clone\_lib="CHORI-501"  
**ORIGIN**  
 Query Match 81.7%; Score 18.9; DB 9; Length 158033;  
 Best Local Similarity 90.9%; Pred. No. 57;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTATGC 22  
 DB 16379 CAGACCATGTGTCAACTATGC 16358  
**RESULT 54**  
**HS1033B10**  
**LOCUS** Human DNA sequence from clone RPS-1033B10 on chromosome 6 open  
**DEFINITION** reading frame 11 (BING4), the B3GALT4 gene for UDP-Gal:betaGlcNAc  
 beta 1,3-galactosyltransferase polypeptide 4, the RPS18 gene for  
 40S ribosomal protein S18, the SACM2L gene for SAC2 (suppressor of  
 actin mutation 2, yeast, homolog)-like (AREL), an HIV Tat specific  
 factor 1 (HTATSP1) pseudogene, a zinc finger pseudogene, the RING1  
 gene for ring finger protein 1 (RNFL), the FABGL gene for FabG  
 (beta-ketoacyl-l-acyl-carrier-protein) reductase, E coli) like  
 protein, the HKE4 gene for HLA class II region protein KE4, the  
 RXRB gene for retinoid X receptor beta, the COL1A2 gene for  
 collagen type XI alpha 2, an HLA class II histocompatibility  
 antigen D or S beta pseudogene and three novel genes. Contains  
 seven CpG islands, ESTs, STSS, and GSSs, complete sequence.  
**ACCESSION** AL031228  
**VERSION** AL031228.1 GI:3646023  
**KEYWORDS** HTG; ARE1; B3GALT4; BING4; C6ORF11; COL1A2; collagen; CpG island;  
 FABGL; HKE4; HTATSP1; retinoid X receptor beta; ribosomal protein;  
 ring finger; RING1; RING2; RNFL; RPS18; RXRB; SACM2L; zinc finger.  
**SOURCE** Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 175737)  
 TUBBY.3.  
 Direct Submission  
 Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonerequests@sanger.ac.uk  
 On Sep 24, 1998 this sequence version replaced gi:3550015.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em.: EMBL; Sw.:  
 SWISSPROT; Tr.: TrEMBL; Wp.: WormPEP; Information on the WormPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
 Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6>  
 RP5-1033B10 is from the library RPCI-5 constructed by the group of  
 Peter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2  
 This sequence is the entire insert of clone RP5-1033B10.  
 FEATURES  
 Location/Qualifiers  
 1..175737  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="RZPD:RPCIP704B101033"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="21.2-21.31"  
 /clone="RP5-1033B10"  
 /clone\_lib="RPCI-5"  
 1..61  
 /note="Alu repeat: matches 2..62 of consensus"  
 63..368  
 /note="AluX repeat: matches 1..299 of consensus"  
 374..403  
 /note="Cheshire repeat: matches 2264..2292 of consensus"  
 404..708  
 /note="AluSg repeat: matches 1..305 of consensus"  
 719..1027  
 /note="AluMb repeat: matches 1..303 of consensus"  
 1028..1220  
 /note="Cheshire repeat: matches 2055..2256 of consensus"  
 1227..1583  
 /note="Cheshire repeat: matches 1..376 of consensus"  
 1593..1671  
 /note="MIR repeat: matches 162..241 of consensus"  
 1962..2265  
 /note="AluSg repeat: matches 1..296 of consensus"  
 2305..2609  
 /note="AluSx repeat: matches 1..307 of consensus"  
 2749..2926  
 /note="AluMb repeat: matches 1..165 of consensus"  
 3009..3297  
 /note="AluSg repeat: matches 1..290 of consensus"  
 4125..4385  
 /note="AluSx repeat: matches 20..294 of consensus"

repeat\_region  
 gene  
 mRNA  
 520..5113  
 /note="MIR repeat: matches 110..214 of consensus"  
 5240..7122  
 /gene="C6ORF11"  
 join(<5240..5553,5705..5799,6368..6463,6617..>7122)  
 /gene="C6ORF11"  
 /product="dJ1033B10.2.2 (chromosome 6 open reading frame  
 11 (BING4), isoform 2)"  
 /note="continued from c1CF0811.6 in Em:297184  
 match: ESTs: Em:AW193506 Em:BE617532"  
 /evidence=not experimental  
 join(<5240..5553,5705..5799,6368..6463,6617..>730,  
 6853..7122)  
 /gene="C6ORF11"  
 /product="dJ1033B10.2.1 (chromosome 6 open reading frame  
 11 (BING4), isoform 1)"  
 /note="continued from c1CF0811.6 in Em:297184  
 match: cDNAs: Em:AK009209  
 match: ESTs: Em:AI204504 Em:BF338950 Em:CNSLT1D01  
 Em:AW192492 Em:BG334317 Em:BG476159 Em:CNSLT1FLA"  
 /evidence=not experimental  
 join(<5240..5553,5705..5799,6368..6463,6617..>7122)  
 /gene="C6ORF11"  
 /codon\_start=2  
 /evidence=not experimental  
 /product="dJ1033B10.2.2 (chromosome 6 open reading frame  
 11 (BING4), isoform 2)"  
 /protein\_id="CAC88173.2"  
 /db\_xref="GI:15799226"  
 /db\_xref="SPTREMBL:O96KL2"  
 /translation="YMATSGLDHOLKIFDLRGTYOPLSTRTPHGAGHLAFSORGLLV  
 AGMDVYNWAGQKASPPSLEQSYLTHRLSGPVHGLQCFPEFVGLVGHGGTSM  
 VPGAGENFQGLSENPRSRKQREWEVKALLEKVPALICLDERALAEVDVILEQG  
 KRQIERLGYDPOKAFQPKPKQKGRSSASLVKRRKVMDEHRVNEHHWHPQLPQ  
 AAPSLCLCFQPATSHVPSFSPVTRTSRAFSSTIRRRPSPRGHLPWD  
 LCAPDSRVAWEQSLPKITCREMSPWNKEVAVWPLPOLGVDSCLLGVGKESDF  
 LD"  
 join(<5240..5553,5705..5799,6368..6463,6617..>730,  
 6853..6951)  
 /gene="C6ORF11"  
 /note="match: proteins: Tr:O15213"  
 /codon\_start=2  
 /evidence=not experimental  
 /product="dJ1033B10.2.1 (chromosome 6 open reading frame  
 11 (BING4), isoform 1)"  
 /protein\_id="CAC38438.2"  
 /db\_xref="GI:15799225"  
 /db\_xref="SPTREMBL:O96KL3"  
 /translation="YMATSGLDHOLKIFDLRGTYOPLSTRTPHGAGHLAFSORGLLV  
 AGMDVYNWAGQKASPPSLEQSYLTHRLSGPVHGLQCFPEFVGLVGHGGTSM  
 VPGAGENFQGLSENPRSRKQREWEVKALLEKVPALICLDERALAEVDVILEQG  
 KRQIERLGYDPOKAFQPKPKQKGRSSASLVKRRKVMDEHRVNEHHWHPQLPQ  
 KEAKAKPTGARPSALDFVR"  
 5838..6062  
 /note="v copies 25 mer 88% conserved"  
 complement(6837..7118)  
 /note="match: STS: Em:G23698"  
 7094..7099  
 /gene="C6ORF11"  
 7111  
 /gene="C6ORF11"  
 7119  
 /gene="C6ORF11"  
 7122  
 /gene="C6ORF11"  
 complement(7409..9031)  
 /gene="B3GALT4"  
 complement(7409..9031)  
 /gene="B3GALT4"  
 /product="dJ1033B10.3 (UDP-Gal:betaGlcNAc beta  
 1,3-galactosyltransferase, polypeptide 4)"  
 /note="match: cDNAs: Em:Y15061 Em:AB003478 Em:AB026730  
 match: ESTs: Em:RA365900 Em:H20531 Em:R40008 Em:AA987754



```

/rpt family="RMER15"
17901..18258
/rpt family="L1"
24317..24911
/rpt family="L1"
26609..26995
/rpt family="MaLR"
27862..28567
/rpt family="L1"
28530..28857
/rpt family="L1"
29522..29618
/rpt family="MBR2_type"
29812..29914
/rpt family="MaLR"
29965..30163
/rpt family="B2"
30171..30285
/rpt family="B2"
31369..32022
/rpt family="L1"
32016..32079
/rpt family="L1"
32087..32540
/rpt family="L1"
32534..32703
/rpt family="L1"
32748..32813
/rpt family="B4"
32815..32950
/rpt family="L1"
33378..33454
/rpt family="MaLR"
33498..33598
/rpt family="MaLR"
33599..33757
/rpt family="B4"
33758..33881
/rpt family="MaLR"
34094..37295
/rpt family="L1"
37626..37855
/rpt family="L1"
37961..38171
/rpt family="MBR2_type"
38688..38926
/rpt family="L1"
39031..39086
/rpt family="L1"
39094..39278
/rpt family="MaLR"
40683..41021
/rpt family="L1"
43245..43439
/rpt family="L1"
43440..43557
/rpt family="L1"
43553..44781
/rpt family="L1"
45200..45641
/rpt family="L1"
45760..46148
/rpt family="MaLR"
46159..46338
/rpt family="L1"
5112..51279
/rpt family="L1"
51410..52169
/rpt family="L1"
52317..52532
/rpt family="ERVK"
52641..52774
/rpt family="ERVK"

```

```

repeat_region 52942..53122
/rpt family="ERVK"
53226..53579
/rpt family="L1"
53606..54034
/rpt family="ERVK"
55404..55455
/rpt family="ERV1"
55531..55987
/rpt family="ERVK"
56039..56168
/rpt family="L1"
56152..56351
/rpt family="L1"
56422..57061
/rpt family="L1"
57128..57319
/rpt family="B2"
59836..59901
/rpt family="ERV1"
60007..60329
/rpt family="L1"
61190..61274
/rpt family="B4"
61656..62412
/rpt family="L1"

```

Query Match 81.7%; Score 18.8; DB 10; Length 201923;

Best Local Similarity 90.9%; Pred. No. 59;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23

Db 112473 GGACCATGTGTCAACTGTATGCC 112494

RESULT 56

AC117573

LOCUS AC117573 239439 bp DNA linear ROD 27-SEP-2003  
Mus musculus chromosome 3, clone RP23-30D1, complete sequence.

DEFINITION AC117573

ACCESSION AC117573.9 GI:34330231

VERSION HTG.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 239439)

AUTHORS Birren, B., Nusbaum, C. and Lander, E.

TITLE Mus musculus chromosome 3, clone RP23-30D1

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 239439)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguski, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, I., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

```

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 239439)
REFERENCE
AUTHORS
Birren,B., Nusbaum,C. and Lander,E.
Direct Submission
Submitted (21-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 239439)
REFERENCE
AUTHORS
Birren,B., Nusbaum,C. and Lander,E.
Direct Submission
Submitted (29-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 239439)
REFERENCE
AUTHORS
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Soukhgater,B., Camarata,J., Chang,J., Choepel,Y.,
Collimore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferrelira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Karat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meidrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,C.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 29, 2003 this sequence version replaced gi:33667201.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L22308
Center clone name: 30_D_1
-----
Location/Qualifiers
1..239439
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP23-30D1"
/clone_lib="RPMI-23 Female Mouse BAC"
misc_feature
1..21390
/clone="wgs_end_extension"
/clone_end="SP6"
repeat_region
1..1059
/rpt_family="Lx"
repeat_region
1139..1188
/rpt_family="TG)n"
repeat_region
1488..1534
/rpt_family="(TTG)n"
repeat_region
2866..2957
/rpt_family="(TCC)n"
repeat_region
complement(3464..3808)
/rpt_family="ORR1C"
complement(3906..4030)
/rpt_family="L1VC4a"
4052..4482
/rpt_family="Lx8"
4723..4742
/rpt_family="AT-rich"
4760..4935
/rpt_family="Lx8"
complement(5078..5516)
/rpt_family="RLTR11A"
5605..5861
/rpt_family="(A)n"
6089..6191
/rpt_family="L1VL2"
6186..6573
/rpt_family="L1F"
6571..6751
/rpt_family="L1"
6745..7114
/rpt_family="L1 MM"
complement(7527..7713)
/rpt_family="L2"
7757..7894
/rpt_family="MTE"
complement(8035..8544)
/rpt_family="RMER4B"
complement(8558..8699)
/rpt_family="ORR1D"
9956..9988
/rpt_family="(CA)n"
9989..10010
/rpt_family="AT-rich"
10067..10095
/rpt_family="(TG)n"
10579..10690
/rpt_family="ID-B"
10691..10724
/rpt_family="B1 MM"
10825..10905
/rpt_family="CT-rich"
10905..10976
/rpt_family="(CA)n"
10977..11001
/rpt_family="AT-rich"
11426..11722
/rpt_family="Lx6"
11852..12323
/rpt_family="RMER5"
12899..12964
/rpt_family="(TC)n"
complement(12935..13038)
/notes="single clone coverage"
complement(13001..13161)
/rpt_family="Lx6"
13807..13836
/rpt_family="AT-rich"
14477..14510
/rpt_family="(GA)n"
14796..14830
/rpt_family="AT-rich"
15208..15270
/rpt_family="G-rich"
complement(15340..15670)
/rpt_family="Lx2"
complement(15945..16158)
/rpt_family="B3"
18138..18163
/rpt_family="AT-rich"
18831..18872
/rpt_family="AT-rich"
19557..19766
/rpt_family="RMER6A"
complement(20240..20358)

```

```

repeat_region      /rpt_family="L1MA4A"
20759..20907
/rpt_family="L1F"
20956..26794
/rpt_family="L1_MM"
21391..21396
/notes="clone boundary
clone end:SP6
site:EcotI"
repeat_region      26795..26840
/rpt_family="AT_rich"
27148..27176
/rpt_family="(TTTA)n"
complement(28717..29125)
/rpt_family="MTC"
complement(29530..35730)
/rpt_family="L1F"
complement(36528..36728)
/rpt_family="L1MA6"

```

Query Match 81.7%; Score 18.8; DB 10; Length 239439;  
 Best Local Similarity 90.9%; Pred. No. 60;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 2 GGACCATGTGCAACTTATGCC 23
|||||
Db 4387 GGACCATGTGCAACTTATGCC 4408
|||||

```

```

RESULT 57
AC125780
LOCUS      150206 bp DNA linear HTG 30-JUN-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 11 clone
Ba0030115, ** SEQUENCING IN PROGRESS **, 10 ordered pieces.
ACCESSION  AC125780
VERSION     AC125780.1 GI:21630424
KEYWORDS    HTG; HTGS PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
REFERENCE   1 (bases 1 to 150206)
AUTHORS     Gupta, V., Bharti, A.K., Raghuvaran, S., Khurana, P., Khurana, J.P. and
TVagi, A.K.
TITLE       Oryza sativa (japonica cultivar-group) chromosome 11 BAC clone
JOURNAL     Ba0030115, sequencing in progress
REFERENCE   2 (bases 1 to 150206)
AUTHORS     Gupta, V., Bharti, A.K., Raghuvaran, S., Khurana, P., Khurana, J.P. and
TVagi, A.K.

```

```

TITLE       Direct Submission
JOURNAL     Submitted (30-JUN-2002) Indian Initiative for Rice Genome
            Sequencing, Department of Plant Molecular Biology, University of
            Delhi South Campus, New Delhi, Delhi 110021, India
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * been provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *
            * 1 5865: contig of 5865 bp in length
            * 5866 5965: gap of unknown length
            * 5966 27869: contig of 21904 bp in length
            * 27870 27969: gap of unknown length
            * 27970 34835: contig of 6866 bp in length
            * 34836 34935: gap of unknown length
            * 34936 59567: contig of 24632 bp in length
            * 59568 61457: gap of unknown length
            * 59668

```

```

* 61458 61557: gap of unknown length
* 61558 96355: contig of 34798 bp in length
* 96356 96455: gap of unknown length
* 96456 97670: contig of 1215 bp in length
* 97671 97770: gap of unknown length
* 97771 119278: contig of 21508 bp in length
* 119279 119378: gap of unknown length
* 119379 140583: contig of 21205 bp in length
* 140584 140684: gap of unknown length
* 140684 150206: contig of 9523 bp in length.

```

```

FEATURES
            Location/Qualifiers
            1..150206
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /sub_species="japonica cultivar-group"
            /db_xref="taxon:39947"
            /chromosome="11"
            /clone="Ba0030115"
            /note="japonica cultivar-group"

```

## ORIGIN

Query Match 80.0%; Score 18.4; DB 2; Length 150206;  
 Best Local Similarity 95.0%; Pred. No. 96;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 GGACCATGTGTCACCTTATG 21
|||||
Db 94962 GGACCATGTGTCACCTTATG 94981
|||||

```

```

RESULT 58
BV062317
LOCUS      830 bp DNA linear STS 31-MAY-2003
DEFINITION S212P6441FD12.T0 CZECHII/E1 Mus musculus STS genomic, sequence
tagged site.
ACCESSION  BV062317
VERSION     BV062317.1 GI:31178112
KEYWORDS    STS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 830)
AUTHORS     Wade, C.M., Kullbøkas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,
Lander, E.S., Lindblad-Toh, K., and Daly, M.J.
TITLE       The mosaic structure of variation in the laboratory mouse genome
JOURNAL     Nature 420 (6915), 574-578 (2002)
MEDLINE     22354584
PUBMED      12466352

```

```

COMMENT     Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kersli@genome.wi.mit.edu
            Primer A: None
            Primer B: None
            STS size: 830
            Protocol:
            WGS-discovery: Paired-end low-coverage whole genome shotgun reads
            were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
            reads were placed uniquely on the MGS3V3 C57BL/6J assembly and SNP
            detection was carried out by SSAHA-SNP. 225,000 reads were
            annotated
            as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
            and the strain from which the particular read came. The validation
            rate for these SNPs was estimated at approximately 98%.
            Location/Qualifiers
            1..830
            /organism="Mus musculus"

```





ROMAN, J., SCHAUER, S., SCHUBACK, R., SEAMAN, S., SEVERY, P., SMITH, C.,  
SPENCER, B., STANGE-THOMANN, N., STOJANOVIC, N., STUBBS, M.,  
TALANAS, J., TESTAYE, S., THEODORE, J., TOPHAM, K., TRAVERS, M.,  
VASSILIEV, H., VENKATARAMAN, V.S., VIEL, R., VO, A., WILSON, B., WU, X.,  
WYMAN, D., YOUNG, G., ZAINOUN, J., ZEMBEK, L., ZIMMER, A. and ZODY, M.  
Direct Submission  
Submitted (10-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 10, 2003 this sequence version replaced gl:28273475.  
All repeats were identified using RepeatMasker:  
Sait, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L25279  
Center clone name: 237\_D\_19  
----- Summary Statistics  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 159539 bases at least Q40  
Consensus quality: 160321 bases at least Q30  
Consensus quality: 160613 bases at least Q20  
Insert size: 158000; agarose-fp  
Insert size: 160800; sum-of-contigs  
Quality coverage: 8.6 in Q20 bases; agarose-fp  
Quality coverage: 8.5 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 46702: contig of 46702 bp in length  
\* 46703 46802: gap of 100 bp  
\* 46803 47418: contig of 616 bp in length  
\* 47419 47518: gap of 100 bp  
\* 47519 50731: contig of 3213 bp in length  
\* 50732 50831: gap of 100 bp  
\* 50832 54526: contig of 3695 bp in length  
\* 54527 54626: gap of 100 bp  
\* 54627 65867: contig of 11241 bp in length  
\* 65868 65967: gap of 100 bp  
\* 65968 78498: contig of 12531 bp in length  
\* 78499 97087: gap of 100 bp  
\* 97088 97188: contig of 18489 bp in length  
\* 97189 97994: gap of 100 bp  
\* 97995 142860: contig of 44866 bp in length  
\* 142861 142961: gap of 100 bp  
\* 142962 161700: contig of 18740 bp in length.

FEATURES  
source  
1. .161700  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone\_lib="RP24-237D19"  
/clone\_lib="RPC1-24 Male Mouse BAC"  
1. .46702  
/note="assembly\_fragment"  
clone\_end:SP6  
vector side:left  
46803-47418  
/note="assembly\_fragment"  
47519. .50731  
misc\_feature  
/note="assembly\_fragment"  
50832-54526  
misc\_feature  
/note="assembly\_fragment"  
54627-65867  
misc\_feature  
/note="assembly\_fragment"  
65968-78498  
misc\_feature  
/note="assembly\_fragment"  
78499-97087  
misc\_feature  
/note="assembly\_fragment"  
97088-97188  
misc\_feature  
/note="assembly\_fragment"  
97189-97994  
misc\_feature  
/note="assembly\_fragment"  
97995-142860  
misc\_feature  
/note="assembly\_fragment"  
142861-142961  
misc\_feature  
/note="assembly\_fragment"  
142962-161700

misc\_feature  
/note="assembly\_fragment"  
50832-54526  
misc\_feature  
/note="assembly\_fragment"  
54527-55867  
misc\_feature  
/note="assembly\_fragment"  
55868-57498  
misc\_feature  
/note="assembly\_fragment"  
57499-59087  
misc\_feature  
/note="assembly\_fragment"  
59088-59789  
misc\_feature  
/note="assembly\_fragment"  
59790-59860  
misc\_feature  
/note="assembly\_fragment"  
59861-61700  
clone\_end:T7  
vector side:right

ORIGIN  
Query Match 79.1%; Score 18.2; DB 2; Length 161700;  
Best Local Similarity 87.0%; Pred. No. 1.3e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 CGGACCATGTGTCAACTTATGCC 23  
DB 71137 CTGTCATGTGTCAACTTATGTC 71115  
Search completed: April 20, 2004, 09:33:01  
Job time : 707.871 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 147.606 Seconds

Perfect score: 23  
Sequence: 1 cggaccatgtgtcaacttatgcc 23  
661.956 Million cell updates/sec

Title: US-09-877-819B-34

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : N\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002s:\*

7: Geneseq2003as:\*

8: Geneseq2003bs:\*

9: Geneseq2003cs:\*

10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	23	100.0	23	6	AAL48198 Human HLA
2	23	100.0	26	6	ABL30954 Human HLA
3	23	100.0	255	6	AAL48219 Human HLA
4	23	100.0	267	4	AAL54418 Human hae
5	23	100.0	267	4	AAL54603 Human hae
6	23	100.0	272	4	AAL54317 Human hae
7	23	100.0	294	4	AAL54555 Human hae
8	23	100.0	294	4	AAL54805 Human hae
9	23	100.0	362	3	AAL43013 Human sec
10	23	100.0	396	2	AAV86130 EST clone
11	23	100.0	410	8	ACH49248 Human leu
12	23	100.0	436	3	AAA443818 Mouse sec
13	23	100.0	466	2	AAV86216 EST clone
14	23	100.0	476	3	AAC00117 Human sec
15	23	100.0	490	8	ACH49928 Human leu
16	23	100.0	576	3	AAC00116 Human sec
17	23	100.0	1202	2	AAQ25060 pSalpha
18	23	100.0	1259	4	AAS31123 Human dia
19	23	100.0	1348	3	AAL18332 Lung canc
20	23	100.0	1466	6	ABK64796 Human ben
21	20	87.0	21	9	AAL28584 Human HLA
22	20	87.0	21	9	AAQ28584 PCR prime
23	19.8	86.1	27	2	AAL54603 HLA-DP DP



```

XX OS Homo sapiens.
XX XX
XX PN WO200194639-A1.
XX XX
XX PD 13-DEC-2001.
XX XX
XX PF 07-JUN-2001; 2001WO-US018590.
XX XX
XX PR 09-JUN-2000; 2000US-0210759P.
XX XX
XX (REGC ) UNIV CALIFORNIA.
XX 2A
XX PI White PS, Torrey DC;
XX XX
XX DR WPI; 2002-566450/60.
XX XX
XX XX
XX PT Identifying sequences useful as address/capture tags for flow cytometry
XX PT based minisequencing, by generating tag sequences and rejecting sequences
XX PT based on certain parameters e.g. sequences which form stable hairpins.
XX XX
XX PS Disclosure; Fig 4; 35pp; English.
XX XX
XX XX The present invention relates to a method of identifying sequences useful
XX CC as address/capture tags, involving rejecting sequences having common sub-
XX CC sequences with a sub-sequence length greater than specified number of
XX CC bases, and sequences which can form stable hairpins and stable dimers
XX CC from a sample of oligonucleotides, and selecting those sequences with a high
XX CC degree of specificity. The method is useful for identifying a set of
XX CC sequences useful as address/capture tags which can be used for
XX CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow
XX CC cytometry assay. The present sequence is a fragment of the human DPAL
XX CC gene described in the exemplification of the invention
XX XX
XX SQ Sequence 255 BP; 67 A; 57 C; 66 G; 65 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGGACCATGTGTCAACTTATGCC 23
Db 8 CGGACCATGTGTCAACTTATGCC 30

RESULT 4
AAK54418
ID AAK54418 standard; cDNA; 267 BP.
XX AC AAK54418;
XX XX
XX DT 13-NOV-2001 (first entry)
XX XX
XX DE Human haematological malignancy-related antigen coding sequence #143.
XX XX
XX KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200164886-A2.
XX XX
XX PD 07-SEP-2001.
XX XX
XX PF 01-MAR-2001; 2001WO-US007272.
XX XX
XX PR 01-MAR-2000; 2000US-0186126P.
XX PR 17-MAR-2000; 2000US-0190479P.
XX PR 27-APR-2000; 2000US-0200545P.
XX PR 28-APR-2000; 2000US-0200303P.
XX PR 28-APR-2000; 2000US-0200779P.

RESULT 3
AAL48219
ID AAL48219 standard; DNA; 255 BP.
XX AC AAL48219;
XX XX
XX DT 01-OCT-2002 (first entry)
XX XX
XX DE Human HLA DPAL exon 2 sequence.
XX XX
XX KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX KW flow cytometry; human; DPAL; gene; ds.

```

PR 01-MAY-2000; 2000US-0200999P.  
PR 04-MAY-2000; 2000US-0202084P.  
PR 22-MAY-2000; 2000US-0206201P.  
PR 14-JUL-2000; 2000US-0218950P.  
PR 03-AUG-2000; 2000US-0222903P.  
PR 04-AUG-2000; 2000US-0223416P.  
PR 07-AUG-2000; 2000US-0223378P.  
XX (CORI-) CORIXA CORP.  
XX  
XX Gaiger A, Algate PA, Mannion J;  
XX WPI; 2001-514842/56.  
XX  
XX Compositions and methods for the detection of hematological malignancies,  
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
XX  
XX Claim 31; Page 371; 1252pp; English.  
XX  
XX The present invention relates to compositions and methods for the  
XX detection, diagnosis and therapy of haematological malignancies. The  
XX present sequence is the coding sequence of a human haematological  
XX malignancy related antigen. The methods of the present invention comprise  
XX detecting the presence of haematological malignancy related antigen(s) in  
XX a sample obtained from the patient (an increased level of the  
XX polypeptide, compared to an unaffected individual, is indicative of an  
XX increased risk). Haematological malignancies which can be treated using  
XX the present invention are chronic lymphocytic leukaemia, lymphoma,  
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
XX cell non-Hodgkin's lymphoma  
XX  
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 23; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.081;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCACACTTATGCC 23  
DB 129 CGGACCATGTGTCACACTTATGCC 151  
RESULT 5  
AAK54603  
ID AAK54603 standard; cDNA; 267 BP.  
XX  
XX AAK54603;  
XX  
XX 13-NOV-2001 (first entry)  
XX  
XX Human haematological malignancy-related antigen coding sequence #328.  
XX  
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;  
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200164886-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 01-MAR-2001; 2001WO-US007272.  
XX  
XX 01-MAR-2000; 2000US-0186126P.  
XX 17-MAR-2000; 2000US-0190479P.  
XX 27-APR-2000; 2000US-0200545P.  
XX 28-APR-2000; 2000US-0200303P.  
XX 28-APR-2000; 2000US-0200779P.  
XX 01-MAY-2000; 2000US-0200599P.  
XX 04-MAY-2000; 2000US-C202084P.  
XX 22-MAY-2000; 2000US-0206201P.  
XX 14-JUL-2000; 2000US-0218950P.  
XX 03-AUG-2000; 2000US-0222903P.  
XX 04-AUG-2000; 2000US-0206201P.

PR 14-JUL-2000; 2000US-0218950P.  
PR 03-AUG-2000; 2000US-0222903P.  
PR 04-AUG-2000; 2000US-0223416P.  
PR 07-AUG-2000; 2000US-0223378P.  
XX (CORI-) CORIXA CORP.  
XX  
XX Gaiger A, Algate PA, Mannion J;  
XX WPI; 2001-514842/56.  
XX  
XX Compositions and methods for the detection of hematological malignancies,  
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
XX  
XX Claim 31; Page 418; 1252pp; English.  
XX  
XX The present invention relates to compositions and methods for the  
XX detection, diagnosis and therapy of haematological malignancies. The  
XX present sequence is the coding sequence of a human haematological  
XX malignancy related antigen. The methods of the present invention comprise  
XX detecting the presence of haematological malignancy related antigen(s) in  
XX a sample obtained from the patient (an increased level of the  
XX polypeptide, compared to an unaffected individual, is indicative of an  
XX increased risk). Haematological malignancies which can be treated using  
XX the present invention are chronic lymphocytic leukaemia, lymphoma,  
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
XX cell non-Hodgkin's lymphoma  
XX  
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 23; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.081;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCACACTTATGCC 23  
DB 129 CGGACCATGTGTCACACTTATGCC 151  
RESULT 6  
AAK54317  
ID AAK54317 standard; cDNA; 272 BP.  
XX  
XX AAK54317;  
XX  
XX 13-NOV-2001 (first entry)  
XX  
XX Human haematological malignancy-related antigen coding sequence #42.  
XX  
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;  
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200164886-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 01-MAR-2001; 2001WO-US007272.  
XX  
XX 01-MAR-2000; 2000US-0186126P.  
XX 17-MAR-2000; 2000US-0190479P.  
XX 27-APR-2000; 2000US-0200545P.  
XX 28-APR-2000; 2000US-0200303P.  
XX 28-APR-2000; 2000US-0200779P.  
XX 01-MAY-2000; 2000US-0200599P.  
XX 04-MAY-2000; 2000US-C202084P.  
XX 22-MAY-2000; 2000US-0206201P.  
XX 14-JUL-2000; 2000US-0218950P.  
XX 03-AUG-2000; 2000US-0222903P.  
XX 04-AUG-2000; 2000US-0223416P.

PR 07-AUG-2000; 2000US-0223378P.  
XX (CORI-) CORIXA CORP.  
XX Gaiger A, Algate PA, Mannion J;  
XX WPI; 2001-514842/56.  
XX Compositions and methods for the detection of hematological malignancies,  
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
XX Claim 31; Page 345; 1252pp; English.  
XX The present invention relates to compositions and methods for the  
XX detection, diagnosis and therapy of hematological malignancies. The  
XX present sequence is the coding sequence of a human haematological  
XX malignancy related antigen. The methods of the present invention comprise  
XX detecting the presence of haematological malignancy related antigen(s) in  
XX a sample obtained from the patient (an increased level of an  
XX polypeptide, compared to an unaffected individual, is indicative of an  
XX increased risk). Haematological malignancies which can be treated using  
XX the present invention are chronic lymphocytic leukaemia, lymphoma,  
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
XX cell non-Hodgkin's lymphoma  
XX Sequence 272 BP; 67 A; 60 C; 75 G; 70 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 23; DB 4; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.081;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
DB 134 CGGACCATGTGTCAACTTATGCC 156  
RESULT 7  
AAK54555/c  
ID AAK54555 standard; cDNA; 294 BP.  
XX AAK54555;  
XX 13-NOV-2001 (first entry)  
XX Human haematological malignancy-related antigen coding sequence #280.  
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;  
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
XX Homo sapiens.  
XX WO200164886-A2.  
XX 07-SEP-2001.  
XX 01-MAR-2001; 2001WO-US007272.  
XX 01-MAR-2000; 2000US-0186126P.  
XX 17-MAR-2000; 2000US-0190479P.  
XX 27-APR-2000; 2000US-0200545P.  
XX 28-APR-2000; 2000US-0200303P.  
XX 28-APR-2000; 2000US-0200779P.  
XX 01-MAY-2000; 2000US-0200999P.  
XX 04-MAY-2000; 2000US-0202084P.  
XX 22-MAY-2000; 2000US-0206201P.  
XX 14-JUL-2000; 2000US-0218950P.  
XX 03-AUG-2000; 2000US-0222903P.  
XX 04-AUG-2000; 2000US-0223416P.  
XX 07-AUG-2000; 2000US-0223378P.  
XX (CORI-) CORIXA CORP.  
XX Gaiger A, Algate PA, Mannion J;  
XX

XX Gaiger A, Algate PA, Mannion J;  
XX WPI; 2001-514842/56.  
XX Compositions and methods for the detection of hematological malignancies,  
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
XX Claim 31; Page 406; 1252pp; English.  
XX The present invention relates to compositions and methods for the  
XX detection, diagnosis and therapy of haematological malignancies. The  
XX present sequence is the coding sequence of a human haematological  
XX malignancy related antigen. The methods of the present invention comprise  
XX detecting the presence of haematological malignancy related antigen(s) in  
XX a sample obtained from the patient (an increased level of the  
XX polypeptide, compared to an unaffected individual, is indicative of an  
XX increased risk). Haematological malignancies which can be treated using  
XX the present invention are chronic lymphocytic leukaemia, lymphoma,  
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
XX cell non-Hodgkin's lymphoma  
XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 23; DB 4; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
DB 139 CGGACCATGTGTCAACTTATGCC 117  
RESULT 8  
AAK54805/c  
ID AAK54805 standard; cDNA; 294 BP.  
XX AAK54805;  
XX 13-NOV-2001 (first entry)  
XX Human haematological malignancy-related antigen coding sequence #530.  
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;  
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
XX Homo sapiens.  
XX WO200164886-A2.  
XX 07-SEP-2001.  
XX 01-MAR-2001; 2001WO-US007272.  
XX 01-MAR-2000; 2000US-0186126P.  
XX 17-MAR-2000; 2000US-0190479P.  
XX 27-APR-2000; 2000US-0200545P.  
XX 28-APR-2000; 2000US-0200303P.  
XX 28-APR-2000; 2000US-0200779P.  
XX 01-MAY-2000; 2000US-0200999P.  
XX 04-MAY-2000; 2000US-0202084P.  
XX 22-MAY-2000; 2000US-0206201P.  
XX 14-JUL-2000; 2000US-0218950P.  
XX 03-AUG-2000; 2000US-0222903P.  
XX 04-AUG-2000; 2000US-0223416P.  
XX 07-AUG-2000; 2000US-0223378P.  
XX (CORI-) CORIXA CORP.  
XX Gaiger A, Algate PA, Mannion J;  
XX

DR WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX  
 XX Claim 31; Page 469; 1252pp; English.  
 XX  
 CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of hematological malignancies. The  
 CC present sequence is the coding sequence of a human hematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of hematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Hematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX  
 XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 23; DB 4; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 0.082;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCACTTATGCC 23  
 DB 139 CGGACCATGTGTCACTTATGCC 117  
 RESULT 9  
 AAA43013  
 ID AAA43013 standard; cDNA; 362 BP.  
 XX  
 XX AAA43013;  
 AC  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Human secreted expressed sequence tag SEQ ID NO:1753.  
 XX  
 XX Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;  
 XX expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;  
 KW antiulcer; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
 KW autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
 XX infection; depression; psoriasis; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200021990-A1.  
 PN  
 XX  
 XX 20-APR-2000.  
 PD  
 XX  
 XX 15-OCT-1999; 99WO-US024205.  
 PF  
 XX  
 XX 15-OCT-1998; 98US-0104435P.  
 PR  
 XX  
 XX (GEMY ) GENETICS INST INC.  
 PA  
 XX  
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M;  
 PI  
 XX WPI; 2000-317937/27.  
 DR  
 XX

PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (sESTs), useful for treating various disorders  
 XX such as autoimmune, infectious, and central nervous system disorders.  
 XX  
 XX Claim 1; Page 526; 618pp; English.  
 XX  
 CC AAA41261 to AAA43419 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, xenopus and rat tissue  
 CC sources. The sESTs can have a range of activities depending on the  
 CC tissues they are isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antiulcer;  
 CC osteoprotective; nootropic; antiparkinsonian; antipsoriatic;  
 CC cerebroprotective; anticonvulsant; and antidepressant. The sESTs are  
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention  
 XX  
 XX Sequence 362 BP; 86 A; 86 C; 99 G; 91 T; 0 U; 0 Other;  
 SQ

Query Match 100.0%; Score 23; DB 3; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.084;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCACTTATGCC 23  
 DB 133 CGGACCATGTGTCACTTATGCC 155  
 RESULT 10  
 AAV86130  
 ID AAV86130 standard; cDNA; 396 BP.  
 XX  
 XX AAV86130;  
 AC  
 XX  
 DT 27-APR-1999 (first entry)  
 XX  
 DE EST clone H45.  
 XX  
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO9845435-A2.  
 PN  
 XX  
 XX 15-OCT-1998.  
 PD  
 XX  
 XX 10-APR-1998; 98WO-US006954.  
 PF  
 XX  
 XX 10-APR-1997; 97US-00835913.  
 PR  
 XX  
 XX (GEMY ) GENETICS INST INC.  
 PA  
 XX  
 XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 PI Spaulding V, Agostino MD;  
 PI  
 XX WPI; 1999-070076/06.  
 DR  
 XX  
 XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT

PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.

XX Claim 1; Page 131; 633pp; English.

XX This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 396 BP; 100 A; 97 C; 101 G; 98 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 2; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.085;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACCTATGCCC 23  
 Db 152 CGGACCATGTGTCACCTATGCCC 174

RESULT 11  
 ACH49248  
 ID ACH49248 standard; cDNA; 410 BP.

XX ACH49248;

DT 13-OCT-2003 (first entry)

DE Human leukocyte cDNA #842.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful  
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
 PT mapping, in the recombinant production of protein, or in generating  
 PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 36450; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 CC determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to  
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
 CC are useful in diagnostics as expressed sequence tags (EST) for  
 CC identifying expressed genes or for physical mapping of the human genome,  
 CC in forensics, in assessing biodiversity, or in identifying mutations  
 CC responsible for genetic disorders and other traits. The nucleotide  
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
 CC for chromosome and gene mapping, in the recombinant production of  
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
 CC is useful for generating antibodies specific for it. The present sequence  
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;

SQ Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 8; Length 410;

Best Local Similarity 100.0%; Pred. No. 0.086;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACCTATGCCC 23

Db 152 CGGACCATGTGTCACCTATGCCC 174

RESULT 12

AAA43818

ID AAA43818 standard; cDNA; 436 BP.

XX AAA43818;

XX 21-AUG-2000 (first entry)

XX Mouse secreted expressed sequence tag SEQ ID NO:393.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cyostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiasthmatic; vulnerrary; antiparkinsonian;  
 KW antitumor; osteopathic; neuroprotective; nontropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
 KW autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
 KW infection; depression; psoriasis; ss.

XX Mus musculus.

XX WO200021991-A1.

XX 20-APR-2000.

XX 15-OCT-1999; 99WO-US024206.

XX 15-OCT-1998; 98US-0104436P.

XX (GBMY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy M, Bowman MR;

XX WPI; 2000-3-7938/27.

XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (sESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders.

XX Claim 1; Page 306; 803pp; English.



XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue  
 CC sources. The sESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antitumor;  
 CC osteoprotective; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;  
 CC cerebroprotective; anticoagulant; and antidepressant. The sESTs can be  
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention

XX Sequence 436 BP; 108 A; 113 C; 106 G; 109 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 3; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 0.086;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23

Db 174 CGGACCATGTGTCAACTTATGCC 196

RESULT 13  
 AAV86216  
 ID AAV86216 standard; cDNA; 456 BP.  
 XX AC AAV86216;  
 XX 27-APR-1999 (first entry)  
 XX EST clone O67.  
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX Homo sapiens.

XX WO9845435-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98WO-US006954.

XX 10-APR-1997; 97US-00835913.

XX (GENY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 XX Spaulding V, Agostino MJ;

XX WPI; 1999-070076/06.

XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 XX placental, retina and colon cDNA libraries.

XX Claim 1; Page 160-161; 633pp; English.

CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity, haemostatic  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, anti-inflammatory  
 CC and thrombolytic activity, receptor/ligand activity, tumour inhibition  
 CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy

XX Sequence 466 BP; 110 A; 122 C; 122 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 2; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 0.087;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23

Db 120 CGGACCATGTGTCAACTTATGCC 142

RESULT 14  
 AAC00117  
 ID AAC00117 standard; cDNA; 476 BP.

XX AC AAC00117;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 115.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00111.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 115; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
 XX mRNAs encoding secreted proteins. An ORF has been identified within the  
 XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 XX derived from 30 different tissues. EST sequences usually correspond  
 XX mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 XX well suited for isolating cDNA sequences derived from the 5' ends of  
 XX mRNAs and even in those cases where longer cDNA sequences have been  
 XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 XX mRNAs with intact 5' ends and can therefore be used to obtain full length  
 XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 XX gene therapy and chromosome mapping procedures. They are used to obtain

```
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 23; DB 3; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 195 CGGACCATGTGTCAACTTATGCC 217

RESULT 15
ACH49928
ID ACH49928 standard; cDNA; 490 BP.
XX AC ACH49928;
XX DT 13-OCT-2003 (first entry)
XX DE Human leukocyte cDNA #1522.
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DNA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX Claim 1; SEQ ID NO 37140; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX Sequence 490 BP; 112 A; 130 C; 125 G; 121 T; 0 U; 2 Other;
```

```
Query Match 100.0%; Score 23; DB 8; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 225 CGGACCATGTGTCAACTTATGCC 247

RESULT 16
AAC00116
ID AAC00116 standard; cDNA; 576 BP.
XX AC AAC00116;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 114.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00110.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID NO 114; 71pp + Sequence Listing; English.
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors
XX Sequence 576 BP; 131 A; 158 C; 141 G; 145 T; 0 U; 1 Other;

Query Match 100.0%; Score 23; DB 3; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 295 CGGACCATGTGTCAACTTATGCC 317

RESULT 17
```

AAQ25060  
 ID AAQ25060 standard; DNA; 1202 BP.  
 AC AAQ25060;  
 XX  
 XX 15-NOV-1992 (first entry)  
 XX  
 XX PSBalpha-318 clone.  
 XX  
 XX  
 XX RFLP; probe p29G8; beta-lymphoblastoid cell line; LG2; T5-1; 6.3.6;  
 KW transplant; transduction; paternity; ss.  
 XX  
 XX Synthetic.  
 XX  
 XX US5110920-A.  
 XX  
 XX 05-MAY-1992.  
 XX  
 XX 05-DEC-1984; 84US-00678255.  
 XX  
 XX 22-JAN-1982; 82US-00341902.  
 PR 07-JAN-1983; 83US-00456373.  
 PR 30-AUG-1988; 88US-00238619.  
 XX  
 XX (CETU ) CETUS CORP.  
 PA  
 XX Erlich HA;  
 XX  
 XX WPI; 1992-175244/21.  
 XX  
 XX New DNA probes specific to single class II HLA locus - useful in HLA  
 PT typing e.g. to evaluate paternity and transplant or transduction  
 PT compatibility and to diagnose disease susceptibility.  
 XX  
 XX Disclousre; Page 11; 21pp; English.  
 XX  
 XX The sequence given is a pBSalpha-318 clone which was derived from a beta-  
 CC lymphoblastoid cell line IG2 cDNA library using a probe designated p29G8.  
 CC This probe bound to sequences distinct from those which lead to the  
 CC elucidation of HLA-Dp34 (see also AAQ25059). A genomic blot pattern with  
 CC DNA from the cell lines T5-1 and its HLA hemizygous derivative 6.3.6  
 CC indicates that the p29G8 locus maps within the HLA region. p29G8 has been  
 CC found to be a HLA-SBalpha clone and could be used to isolate the given  
 CC sequence. p29G8 could be useful in HLA typing based on RFLPs. It can be  
 CC utilized in paternity disputes or for determining transplant or transduction  
 CC compatibility. It can also be used to make disease correlations to  
 CC diagnose diseases or predict susceptibility to diseases .  
 XX  
 XX Sequence 1202 BP; 289 A; 336 C; 277 G; 300 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 23; D3 2; Length 1202;  
 Best Local Similarity 100.0%; Pred. No. 0.099;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACCATGTGTCAACTTATGCC 23  
 D5 180 CGGACCATGTGTCAACTTATGCC 202  
 RESULT 18  
 AAAS31123  
 ID AAAS31123 standard; cDNA; 1259 BP.  
 XX  
 XX AAAS31123;  
 AC  
 XX 04-DEC-2001 (first entry)  
 DT  
 XX Human diagnostic and therapeutic polynucleotide (DITHP) #138.  
 DE  
 XX Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
 KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
 KW respiratory disorder; ss.

XX  
 OS Homo sapiens.  
 XX  
 XX WO200162927-A2.  
 XX  
 XX 30-AUG-2001.  
 XX  
 XX 21-FEB-2001; 2001WO-US006059.  
 XX  
 XX 24-FEB-2000; 2000US-0184693P.  
 PR 24-FEB-2000; 2000US-0184697P.  
 PR 24-FEB-2000; 2000US-0184698P.  
 PR 24-FEB-2000; 2000US-0184768P.  
 PR 24-FEB-2000; 2000US-0184769P.  
 PR 24-FEB-2000; 2000US-0184770P.  
 PR 24-FEB-2000; 2000US-0184771P.  
 PR 24-FEB-2000; 2000US-0184772P.  
 PR 24-FEB-2000; 2000US-0184773P.  
 PR 24-FEB-2000; 2000US-0184774P.  
 PR 24-FEB-2000; 2000US-0184776P.  
 PR 24-FEB-2000; 2000US-0184777P.  
 PR 24-FEB-2000; 2000US-0184797P.  
 PR 24-FEB-2000; 2000US-0184813P.  
 PR 24-FEB-2000; 2000US-0184837P.  
 PR 24-FEB-2000; 2000US-0184841P.  
 PR 24-FEB-2000; 2000US-0185213P.  
 PR 24-FEB-2000; 2000US-0185216P.  
 PR 12-MAY-2000; 2000US-0203785P.  
 PR 15-MAY-2000; 2000US-0204228P.  
 PR 16-MAY-2000; 2000US-0204252P.  
 PR 16-MAY-2000; 2000US-0204821P.  
 PR 16-MAY-2000; 2000US-0204908P.  
 PR 16-MAY-2000; 2000US-0205232P.  
 PR 17-MAY-2000; 2000US-0204815P.  
 PR 17-MAY-2000; 2000US-0204863P.  
 PR 17-MAY-2000; 2000US-0205221P.  
 PR 17-MAY-2000; 2000US-0205285P.  
 PR 17-MAY-2000; 2000US-0205286P.  
 PR 17-MAY-2000; 2000US-0205287P.  
 PR 17-MAY-2000; 2000US-0205323P.  
 PR 17-MAY-2000; 2000US-0205324P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
 PI Chen A, D'sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE, Dufour CE;  
 PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TP;  
 PI Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;  
 PI Wright RJ, Yap PB, Yu JY, Bradley DL, Bratcher SR, Chen W;  
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
 XX WPI; 2001-502867/55.  
 XX P-PSDB; AAU19552.  
 DR  
 XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 XX  
 XX Claim 1; Page 361; 522pp; English.  
 XX  
 XX The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DITHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide  
 CC expression by re-codifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DITHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and

Wed Apr 21 10:46:40 2004

us-09-877-819b-34.rng

CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DTHPs and in assays to identify modulators of DTHP  
 CC expression and activity. The anti-DTHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DTHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DTHPs in samples (e.g. by enzyme linked immunosorbant assay  
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic  
 CC (DTHP) polynucleotides of the invention

XX Sequence 1259 BP; 293 A; 350 C; 300 G; 316 T; 0 U; 0 Other;

Query Match 100.0%; Score 23; DB 4; Length 1259;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 Db 315 CGGACCATGTGTCAACTTATGCC 337

RESULT 19  
 AAF18332  
 ID AAF18332 standard; DNA; 1348 BP.

XX AC AAF18332;

XX DT 14-MAR-2001 (first entry)

XX DE Lung cancer associated polynucleotide sequence SEQ ID 351.

XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.

XX OS Homo sapiens.

XX PN WO200055180-A2.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US005918.

XX PR 12-MAR-1999; 99US-0124270P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (ROSE/) ROSEN C A.

XX PI Ruben SM;

XX PS WPI; 2000-587514/55.

XX DR P-PSDB; AAB58456.

XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer.

XX PS Claim 1; Page 808-809; 1425pp; English.

XX CC Polynucleotide sequences AAF17962 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective, cytostatic, cardioactive;  
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the protein  
 CC or polynucleotide sequences. The lung cancer associated polynucleotide  
 CC sequences may be used for detection of lung cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The proteins may be used to treat disorders such as  
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary, also  
 CC cardiovascular, renal, and proliferative disorders. The proteins may also  
 CC be used in the treatment of wounds and infectious diseases.

CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are  
 CC used in the course of the invention for the identification and  
 CC characterisation of the polynucleotide and protein sequences

XX Sequence 1348 BP; 335 A; 373 C; 307 G; 324 T; 0 U; 9 Other;

Query Match 100.0%; Score 23; DB 3; Length 1348;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 Db 253 CGGACCATGTGTCAACTTATGCC 275

RESULT 20

ABK64796

ID ABK64796 standard; DNA; 14646 BP.

XX AC ABK64796;

XX DT 18-JUN-2002 (first entry)

XX DE Human benign prostatic hyperplasia gene #691.

XX KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

XX OS Homo sapiens.

XX PN WO200212440-A2.

XX PD 14-FEB-2002.

XX PF 07-AUG-2001; 2001WO-US024708.

XX PR 07-AUG-2000; 2000US-0223323P.

XX PR 05-JUN-2001; 2001US-00873319.

XX PA (GENE-) GENE LOGIC INC.

XX PA (NISE) JAPAN TOBACCO INC.

XX PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;

XX WPI; 2002-257476/30.

XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 PT detecting expression levels of one or more genes in prostate cells from  
 PT patient that are differentially regulated compared to normal prostate  
 PT cells.

XX Disclosure; Page 373-377; 444pp; English.

XX The invention relates to a method of diagnosing (I) the onset or  
 CC progression of benign prostatic hyperplasia (BPH); or screening (II) for  
 CC or identifying an agent that modulates the onset or progression of BPH.  
 CC The method is based on changes in gene expression in BPH tissue isolated  
 CC from patients exhibiting different clinical states of prostate  
 CC hyperplasia as compared to normal prostate tissue. (I) comprises  
 CC detecting the expression levels of one or more genes in prostate cells  
 CC from the subject that are differentially regulated compared to normal  
 CC prostate cells. (II) comprises preparing a first gene expression profile  
 CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles. (I)  
 CC is useful for diagnosing the onset or progression of BPH. (II) is useful  
 CC for identifying an agent that modulates the onset or progression of BPH.  
 CC The methods are useful to present information identifying the expression  
 CC level in a tissue or cells, by comparing the expression level of genes  
 CC given in the specification in the tissue or cells to the level of

CC expression of gene in the database, and displaying the expression levels  
 CC of at least one gene in the tissue or cell sample compared to the  
 CC expression level in BPH. Agents using (II) are useful for treating BPH or  
 CC prostate cancer. ABK64106-ABK6496C represent human benign prostatic  
 CC hyperplasia gene sequences of the invention  
 XX  
 SQ Sequence 14646 BP; 4552 A; 2957 C; 2741 G; 4395 T; 0 U; 1 Other;

Query Match 100.0%; Score 23; DB 6; Length 14646;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
 |||||  
 DB 6334 CGGACCATGTGTCAACTTATGCC 6356

RESULT 21  
 AAQ28584  
 ID AAQ28584 standard; DNA; 21 BP.  
 XX  
 AC AAQ28584;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-FEB-1993 (first entry)  
 XX  
 XX Human HLA DPA1 primer #467.  
 XX  
 KW Biotin; primer; [35S]-alpha-thio-dNTP; radioactive label;  
 KW primer extension; template-directed; sequence-specific labelling;  
 KW DNA polymerase-catalysed extension; primer-template complex; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9215712-A1.  
 XX  
 PD 17-SEP-1992.  
 XX  
 PF 04-MAR-1992; 92WO-US001905.  
 XX  
 PR 05-MAR-1991; 91US-00664837.  
 PR 11-OCT-1991; 91US-00775786.  
 XX  
 PA (MOLE-) MOLECULAR TOOL INC.

PI Goeliet P, Knapp MR, Anderson S;  
 XX  
 DR WPI; 1992-331756/40.  
 XX  
 PT Nucleic acid template-dependent, primer extension reaction - using at  
 PT least two different sequence terminators, for genetic typing.  
 XX  
 PS Disclosure; Page 46; 78pp; English.  
 XX

CC The sequences given in AAQ28578-98 are primers which were used to  
 CC illustrate the methods of the invention. The methods use a reagent  
 CC composition comprising an aqueous carrier and an admixture of at least  
 CC two different terminators of a nucleic acid template-dependent, primer  
 CC extension reaction. Each terminator is capable of specifically  
 CC terminating the reaction in a manner strictly dependant in the identity  
 CC of the unpaired nucleotide base immediately adjacent to, and downstream of  
 CC the 3' end of the primer. At least one of the terminators is labelled  
 CC with a detectable marker eg. 35S. The methods allow analyses of nucleic  
 CC acid sequences that can be useful in the diagnosis of infectious diseases,  
 CC genetic disorders and the identification of individuals and their  
 CC parentage. See also AAQ28575-77. [Updated on 25-MAR-2003 to correct PN  
 CC field.]  
 XX

SQ Sequence 21 BP; 5 A; 5 C; 5 G; 6 T; 0 U; 0 Other;

Query Match 87.0%; Score 20; DB 2; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTAT 23  
 |||||  
 DB 2 CGGACCATGTGTCAACTTAT 21

RESULT 22  
 ADE10080  
 ID ADE10080 standard; DNA; 21 BP.  
 XX  
 AC ADE10080;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE PCR primer #1 for DPA exon 2 in human transgenes in pig.

XX  
 KW Transgenic; swine; pig; human leukocyte antigen; HLA-DQ; transgene;  
 KW xenogenic cellular response; organ; tissue; xenotransplantation; human;  
 KW HLA-DR; DPA; PCR; primer; ss.  
 XX  
 OS Homo sapiens.

XX  
 PN US639122-B1.  
 XX  
 PD 28-OCT-2003.  
 XX  
 PF 19-SEP-2000; 2000US-00664846.

XX  
 PR 19-SEP-2000; 2000US-00664846.  
 XX  
 PA (ANIM-) ANIMAL TECHNOLOGY INST TAIWAN.

XX  
 PI Tu C, Lee C, Lee J, Tsuji K;  
 XX  
 DR WPI; 2003-842793/78.

XX  
 PT New transgenic swine whose genome includes a human leukocyte antigen  
 PT (HLA)-DQ transgene, useful for providing organs or tissues for  
 PT xenotransplantation into humans.  
 XX  
 PS Example; Col 5; 11pp; English.

XX  
 CC The present invention relates to a transgenic swine whose genome includes  
 CC a human leukocyte antigen (HLA)-DQ transgene. The HLA-DQ transgene  
 CC comprises a combination of human HLA-DQA1 and HLA-DQB1 transgene operably  
 CC linked to a constitutive promoter, where the expression of the transgene  
 CC reduces a xenogenic cellular response between cells from the transgenic  
 CC swine and a human subject. The transgenic swine and swine cells having  
 CC human HLA-D genes are useful for providing organs or tissues for  
 CC xenotransplantation into humans. The present sequence represents a PCR  
 CC primer used in the examples of the present invention.

SQ Sequence 21 BP; 5 A; 5 C; 5 G; 6 T; 0 U; 0 Other;

Query Match 87.0%; Score 20; DB 9; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTAT 20  
 |||||  
 DB 2 CGGACCATGTGTCAACTTAT 21

RESULT 23  
 AAQ54603  
 ID AAQ54603 standard; DNA; 27 BP.

XX AAQ54603;

XX  
 DT 25-MAR-2003 (revised)  
 DT 29-JUN-1994 (first entry)  
 XX

DE HLA-DP DPalpha genotype determination primer GH98.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 ; Search time 31.2529 Seconds  
(without alignments)  
408.405 Million cell updates/sec

Title: US-09-877-819B-34

Perfect score: 23  
Sequence: 1 cggaccatgtcgaactatgcc 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA:\*

- 1: /cgm2\_6/ptodata/2/ina/5A COMB.seq:\*
- 2: /cgm2\_6/ptodata/2/ina/5B COMB.seq:\*
- 3: /cgm2\_6/ptodata/2/ina/6A COMB.seq:\*
- 4: /cgm2\_6/ptodata/2/ina/6B COMB.seq:\*
- 5: /cgm2\_6/ptodata/2/ina/PCTUS COMB.seq:\*
- 6: /cgm2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	20	87.0	21	3	US-07-775-786-13
2	20	87.0	21	4	US-09-258-133-13
3	20	87.0	21	4	US-09-664-846A-7
4	16.2	70.4	236	4	US-09-313-294A-4368
5	15.6	67.8	453	4	US-09-501-115-23
6	15.6	67.8	1142	4	US-08-936-168A-92
7	15.6	67.8	4392	4	US-09-657-453A-17
8	15.6	67.8	4566	4	US-09-657-453A-3
9	15.6	67.8	4403765	3	US-09-103-840A-2
10	15.6	67.8	4411529	3	US-09-103-840A-1
11	15.4	67.0	2722	4	US-09-390-234-13
12	15.4	67.0	2722	4	US-09-603-311-13
13	15.2	66.1	33	4	US-09-442-021-5
14	15.2	66.1	480	4	US-09-621-976-1327
15	15.2	66.1	610	4	US-09-702-705-1094
16	15.2	66.1	610	4	US-09-736-457-1094
17	15.2	66.1	610	4	US-09-614-124B-1094
18	15.2	66.1	610	4	US-09-671-325-1094
19	15.2	66.1	1323	1	US-09-328-352-1562
20	15.2	66.1	7451	1	US-08-684-672-23
21	15.2	66.1	9465	4	US-09-206-551-1
22	15.2	66.1	66804	4	US-09-740-041-3
23	15.2	66.1	169998	4	US-09-676-610B-24
24	15.2	66.1	197496	4	US-09-877-177A-10
25	15	65.2	370	2	US-08-569-150A-1
26	15	65.2	375	4	US-09-107-532A-1682
27	14.8	64.3	632	3	US-08-998-416-51

1049	64.3	14.8	3	US-08-682-643-1	Sequence 1, Appli
1054	64.3	14.8	2	US-08-566-096A-5	Sequence 5, Appli
1054	64.3	14.8	2	US-08-668-650B-5	Sequence 5, Appli
1054	64.3	14.8	4	US-09-200-673-5	Sequence 5, Appli
1054	64.3	14.8	4	US-09-194-895-5	Sequence 5, Appli
1054	64.3	14.8	4	US-09-447-907-5	Sequence 5, Appli
1054	64.3	14.8	5	PCT-US95-15646-5	Sequence 5, Appli
1374	64.3	14.8	2	US-09-107-532A-3613	Sequence 3613, Ap
1479	64.3	14.8	2	US-08-668-650B-13	Sequence 13, Appl
1479	64.3	14.8	4	US-09-194-895-13	Sequence 13, Appl
1479	64.3	14.8	4	US-09-447-907-13	Sequence 13, Appl
7143	64.3	14.8	3	US-09-381-862-4	Sequence 4, Appli
8106	64.3	14.8	3	US-09-135-241-1	Sequence 1, Appli
276	63.5	14.6	4	US-09-540-236-306	Sequence 306, App
475	63.5	14.6	3	US-09-115-954-9	Sequence 9, Appli
491	63.5	14.6	4	US-09-621-976-14621	Sequence 14621, A
900	63.5	14.6	4	US-09-435-050A-230	Sequence 230, App
900	63.5	14.6	1	US-08-459-264-1	Sequence 1, Appli
900	63.5	14.6	1	US-08-459-263-1	Sequence 1, Appli
1179	63.5	14.6	4	US-09-134-000C-3246	Sequence 3246, Ap
1233	63.5	14.6	4	US-09-540-236-1920	Sequence 1920, Ap
1551	63.5	14.6	4	US-09-625-972-19	Sequence 19, Appl
1742	63.5	14.6	1	US-08-527-227A-6	Sequence 6, Appli
2192	63.5	14.6	1	US-08-035-392-1	Sequence 1, Appli
2192	63.5	14.6	1	US-08-504-511A-1	Sequence 1, Appli
2205	63.5	14.6	1	US-08-035-392-3	Sequence 3, Appli
2205	63.5	14.6	1	US-08-504-511A-3	Sequence 3, Appli
2880	63.5	14.6	3	US-09-115-954-3	Sequence 3, Appli
3842	63.5	14.6	3	US-09-115-954-7	Sequence 7, Appli
3912	63.5	14.6	3	US-09-115-954-1	Sequence 1, Appli
5503	63.5	14.6	2	US-08-726-012B-1	Sequence 1, Appli
5503	63.5	14.6	4	US-09-023-655-989	Sequence 989, App
8439	63.5	14.6	4	US-09-221-017B-473	Sequence 473, App
2641	63.5	14.6	4	US-09-625-972-22	Sequence 22, Appl
48974	63.5	14.6	3	US-08-920-422-17	Sequence 17, Appl
1230025	62.6	14.4	4	US-09-198-452A-1	Sequence 1, Appli
472	62.6	14.4	4	US-09-569-098A-96	Sequence 96, Appl
2574	62.6	14.4	2	US-08-659-251-13	Sequence 13, Appl
2574	62.6	14.4	2	US-09-256-490-13	Sequence 13, Appl
2574	62.6	14.4	5	PCT-US96-11445-13	Sequence 13, Appl
9633	62.6	14.4	6	5223423-1	Patent No. 5223423
10163	62.6	14.4	2	US-08-659-251-1	Sequence 1, Appli
10163	62.6	14.4	3	US-09-256-490-1	Sequence 1, Appli
10163	62.6	14.4	5	PCT-US96-11445-1	Sequence 1, Appli
584	61.7	14.2	4	US-09-621-976-1560	Sequence 1560, Ap
746	61.7	14.2	4	US-09-356-806-2	Sequence 2, Appli
799	61.7	14.2	4	US-09-277-303-1	Sequence 1, Appli
799	61.7	14.2	4	US-09-277-303-3	Sequence 3, Appli
850	61.7	14.2	4	US-09-342-681C-8	Sequence 8, Appli
977	61.7	14.2	4	US-09-671-317-381	Sequence 381, App
1001	61.7	14.2	4	US-09-671-317-382	Sequence 382, App
1062	61.7	14.2	4	US-09-543-681A-1164	Sequence 1164, Ap
1194	61.7	14.2	4	US-08-860-368B-11	Sequence 11, Appl
1194	61.7	14.2	3	US-08-860-368B-12	Sequence 12, Appl
1434	61.7	14.2	4	US-09-489-039A-2415	Sequence 2415, Ap
1470	61.7	14.2	4	US-09-292-225-40	Sequence 40, Appl
1470	61.7	14.2	4	US-09-292-225-42	Sequence 42, Appl
1527	61.7	14.2	4	US-09-292-225-37	Sequence 37, Appl
1527	61.7	14.2	4	US-09-292-225-39	Sequence 39, Appl
1608	61.7	14.2	4	US-09-292-225-20	Sequence 20, Appl
1608	61.7	14.2	4	US-09-292-225-22	Sequence 22, Appl
1621	61.7	14.2	4	US-09-292-225-34	Sequence 34, Appl
1621	61.7	14.2	4	US-09-292-225-36	Sequence 36, Appl
1626	61.7	14.2	4	US-09-134-000C-2015	Sequence 2015, Ap
1665	61.7	14.2	4	US-09-292-225-17	Sequence 17, Appl
1665	61.7	14.2	4	US-09-292-225-19	Sequence 19, Appl
1752	61.7	14.2	4	US-09-292-225-14	Sequence 14, Appl
1752	61.7	14.2	4	US-09-292-225-16	Sequence 16, Appl
1776	61.7	14.2	4	US-09-134-001C-391	Sequence 391, App
2092	61.7	14.2	4	US-09-356-806-7	Sequence 7, Appli
2298	61.7	14.2	4	US-09-543-681A-115	Sequence 115, App
2438	61.7	14.2	4	US-09-302-769-43	Sequence 43, Appl
2576	61.7	14.2	4	US-09-620-312D-1087	Sequence 1087, Ap

```

101 14.2 61.7 2730 1 US-08-344-536-1 Sequence 1, Appli
102 14.2 61.7 2730 3 US-08-920-562-1 Sequence 1, Appli
103 14.2 61.7 3111 4 US-09-543-681A-3564 Sequence 364, Ap
104 14.2 61.7 3318 4 US-09-540-236-1375 Sequence 1379, Ap
105 14.2 61.7 4884 4 US-09-328-352-2478 Sequence 2478, Ap
106 14.2 61.7 5159 2 US-08-146-930-3 Sequence 3, Appli
107 14.2 61.7 5159 3 US-08-458-240-3 Sequence 3, Appli
108 14.2 61.7 5159 5 PCT-US93-03953-3 Sequence 3, Appli
109 14.2 61.7 5708 4 US-09-566-921-21 Sequence 21, Appli
110 14.2 61.7 9542 3 US-08-968-685A-9 Sequence 9, Appli
111 14.2 61.7 28598 4 US-09-593-995-10 Sequence 10, Appli
112 14.2 61.7 62804 4 US-09-800-960-3 Sequence 3, Appli
113 14.2 61.7 62804 4 US-10-096-960-3 Sequence 3, Appli
114 14.2 61.7 99629 4 US-09-596-002-37 Sequence 37, Appli
115 14.2 61.7 111282 4 US-09-754-250-3 Sequence 3, Appli
116 14 60.9 70 1 US-07-832-905B-38 Sequence 38, Appli
117 14 60.9 70 2 US-08-700-787-38 Sequence 38, Appli
118 14 60.9 74 1 US-07-832-905B-37 Sequence 37, Appli
119 14 60.9 74 2 US-08-700-757-37 Sequence 37, Appli
120 14 60.9 100 4 US-08-956-171E-2046 Sequence 2046, Ap
121 14 60.9 258 2 US-08-611-757-103 Sequence 103, App
122 14 60.9 258 5 PCT-US95-05980-103 Sequence 103, App
123 14 60.9 261 4 US-09-313-294A-3286 Sequence 3286, Ap
124 14 60.9 280 4 US-09-313-294A-1472 Sequence 1472, Ap
125 14 60.9 287 4 US-09-313-294A-4492 Sequence 4492, Ap
126 14 60.9 364 4 US-09-313-294A-4861 Sequence 4861, Ap
127 14 60.9 438 4 US-09-543-681A-351 Sequence 351, App
128 14 60.9 445 4 US-09-464-535-1 Sequence 1, Appli
129 14 60.9 538 3 US-08-979-586-3 Sequence 3, Appli
130 14 60.9 538 4 US-09-577-640-3 Sequence 3, Appli
131 14 60.9 843 4 US-09-252-991A-2744 Sequence 2744, Ap
132 14 60.9 895 4 US-09-464-535-21 Sequence 21, Appli
133 14 60.9 906 4 US-09-501-115-15 Sequence 15, Appli
134 14 60.9 1033 1 US-08-592-126-92 Sequence 92, Appli
135 14 60.9 1033 4 US-09-168-595-92 Sequence 92, Appli
136 14 60.9 1251 4 US-09-540-236-733 Sequence 733, App
137 14 60.9 1340 4 US-08-961-527-330 Sequence 330, App
138 14 60.9 1350 4 US-09-107-532A-1812 Sequence 1812, Ap
139 14 60.9 1359 3 US-09-387-574-11 Sequence 11, Appli
140 14 60.9 1359 4 US-09-668-096-11 Sequence 11, Appli
141 14 60.9 1467 1 US-07-881-075-50 Sequence 50, Appli
142 14 60.9 1467 1 US-08-120-827-50 Sequence 50, Appli
143 14 60.9 1467 1 US-08-478-675-50 Sequence 50, Appli
144 14 60.9 1469 1 US-08-461-027-1 Sequence 1, Appli
145 14 60.9 1469 1 US-09-082-501-1 Sequence 1, Appli
146 14 60.9 1469 5 PCT-US92-01836-1 Sequence 1, Appli
147 14 60.9 1505 3 US-08-875-847B-1 Sequence 1, Appli
148 14 60.9 1505 4 US-09-378-842-1 Sequence 1, Appli
149 14 60.9 1573 4 US-09-620-312D-927 Sequence 927, App
150 14 60.9 1623 4 US-09-107-532A-1660 Sequence 1660, Ap
151 14 60.9 1710 4 US-09-489-039A-5338 Sequence 5338, Ap
152 14 60.9 1777 4 US-08-828-199A-1 Sequence 1, Appli
153 14 60.9 1971 4 US-09-347-878-23 Sequence 23, Appli
154 14 60.9 2187 3 US-09-318-448-1 Sequence 1, Appli
155 14 60.9 2196 4 US-09-962-665-15 Sequence 15, Appli
156 14 60.9 2196 4 US-09-963-333-15 Sequence 15, Appli
157 14 60.9 2219 3 US-08-738-000-3 Sequence 3, Appli
158 14 60.9 2219 3 US-09-258-928-3 Sequence 3, Appli
159 14 60.9 2219 4 US-09-660-872A-3 Sequence 3, Appli
160 14 60.9 2219 4 US-09-592-595A-3 Sequence 3, Appli
161 14 60.9 2220 3 US-08-738-000-1 Sequence 1, Appli
162 14 60.9 2220 3 US-09-258-928-1 Sequence 1, Appli
163 14 60.9 2220 4 US-09-660-872A-1 Sequence 1, Appli
164 14 60.9 2220 4 US-09-592-595A-1 Sequence 1, Appli
165 14 60.9 2625 4 US-10-116-370-1 Sequence 1, Appli
166 14 60.9 3605 3 US-09-098-901-1 Sequence 1, Appli
167 14 60.9 3641 1 US-08-030-096-5 Sequence 5, Appli
168 14 60.9 4403 2 US-08-284-941-1 Sequence 1, Appli
169 14 60.9 4403 2 US-08-447-642-1 Sequence 1, Appli
170 14 60.9 4403 3 US-09-236-503-1 Sequence 1, Appli
171 14 60.9 4403 5 PCT-US93-02147A-1 Sequence 1, Appli
172 14 60.9 5092 3 US-09-412-545-1 Sequence 1, Appli
173 14 60.9 16063 4 US-09-801-052-3 Sequence 3, Appli

```

## RESULT 1

US-07-775-786-13

; Sequence 13, Application US/07775786

; Patent No. 6004744

; GENERAL INFORMATION:

; APPLICANT: GOELET, Philip

; APPLICANT: KNAPP, Michael R.

; APPLICANT: ANDERSON, Stephen

; TITLE OF INVENTION: NUCLEIC ACID TYPING BY POLYMERASE

; TITLE OF INVENTION: EXTENSION OF OLIGONUCLEOTIDES USING TERMINATOR MIXTURES

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HALE and DORR

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07775,786

; FILING DATE: 11-OCT-1991

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/664,837

; FILING DATE: 05-MAR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BAKER, Hollie L.

; REGISTRATION NUMBER: 31,321

; REFERENCE/DOCKET NUMBER: 1019.67.115CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)942-8400

; TELEFAX: (202)942-8484

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; PUBLICATION INFORMATION:

; DOCUMENT NUMBER: WO PCT/US92/01905

; FILING DATE: 04-MAR-1992

; PUBLICATION DATE: 17-SEP-1992

Query Match 87.0%; Score 20; DB 3; Length 21;

Best Local Similarity 100.0%; Pred.No. 0.36;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACCTTAT 20

Db 2 CGGACCATGTGTCACCTTAT 21

## RESULT 2

US-09-258-133-13

; Sequence 13, Application US/09258133  
; Patent No. 6537748  
; GENERAL INFORMATION:  
; APPLICANT: Golet, Philip  
; APPLICANT: Knapp, Michael  
; APPLICANT: Anderson, Stephen  
; TITLE OF INVENTION: REAGENT FOR NUCLEIC ACID TYPING BY POLYMERASE EXTENSION  
; FILE REFERENCE: 13017-2  
; CURRENT APPLICATION NUMBER: US/09/258,133  
; CURRENT FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 07/775,786  
; PRIOR FILING DATE: 1991-10-11  
; PRIOR APPLICATION NUMBER: 07/664,837  
; PRIOR FILING DATE: 1991-03-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic primer for human HLA DPA1  
US-09-258-133-13

Query Match 87.0%; Score 20; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACCTTAT 20  
|||||  
DB 2 CGGACCATGTGTCACCTTAT 21

## RESULT 3

US-09-664-846A-7  
; Sequence 7, Application US/09664846A  
; Patent No. 6639122  
; GENERAL INFORMATION:  
; APPLICANT: Tu, Chin-Fu  
; APPLICANT: Tsuji, Kimiyoshi  
; APPLICANT: Lee, Jang-Ming  
; APPLICANT: Lee, Chun-Jean  
; TITLE OF INVENTION: Transgenic Swine Having HLA-D Gene, Swine Cells Thereof and Xenografts  
; FILE REFERENCE: P1199  
; CURRENT APPLICATION NUMBER: US/09/664,846A  
; CURRENT FILING DATE: 2000-09-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Transgenic Swine Cell DNA  
US-09-664-846A-7

Query Match 87.0%; Score 20; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACCTTAT 20  
|||||  
DB 2 CGGACCATGTGTCACCTTAT 21

## RESULT 4

US-09-313-294A-4368  
; Sequence 4368, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalugudi, Raghurath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherma, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 4368  
; LENGTH: 236  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700348308H1  
; NAME/KEY: unsure  
; LOCATION: 206, 232, 234  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-4368

Query Match 70.4%; Score 16.2; DB 4; Length 236;  
Best Local Similarity 85.7%; Pred. No. 40;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCACCTTATG 21  
|||||  
DB 90 CGGACCATGTGTCACCTTAAAG 110

## RESULT 5

US-09-501-115-23  
; Sequence 23, Application US/09501115  
; Patent No. 6552249  
; GENERAL INFORMATION:  
; APPLICANT: Caboon, Rebecca E.  
; APPLICANT: Fader, Gary M.  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs  
; FILE REFERENCE: BB1328 US NA  
; CURRENT APPLICATION NUMBER: US/09/501,115  
; CURRENT FILING DATE: 2000-02-09  
; EARLIER APPLICATION NUMBER: 60/119,585  
; EARLIER FILING DATE: 1999-February-10  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 23  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-501-115-23

Query Match 67.8%; Score 15.6; DB 4; Length 453;  
Best Local Similarity 81.8%; Pred. No. 86;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCACCTTATGCC 23  
|||||  
DB 117 GGGACATATGTCAACTCATGCC 138

## RESULT 6

US-08-936-165A-92  
; Sequence 92, Application US/08936165A  
; Patent No. 6348582  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Lonetto, Michael  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Richard  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Ward, Judith  
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,



/ TITLE OF INVENTION: Polypeptides and Their Uses  
/ NUMBER OF SEQUENCES: 534  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: SmithKline Beecham Corporation  
/ STREET: 709 Swedeland Road  
/ CITY: King of Prussia  
/ STATE: PA  
/ COUNTRY: USA  
/ ZIP: 19406-0939  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette  
/ OPERATING SYSTEM: DOS  
/ SOFTWARE: FastSeq for Windows Version 2.0  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/936,165A  
/ FILING DATE: 24-SEP-1997  
/ CLASSIFICATION: 536  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 60/027,032  
/ FILING DATE: 24-SEP-1996  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Gimmi, Edward R.  
/ REGISTRATION NUMBER: 38,891  
/ REFERENCE/DOCKET NUMBER: P50549  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 610-270-4478  
/ TELEFAX: 610-270-5090  
/ TELEX:  
/ INFORMATION FOR SEQ ID NO: 92:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1142 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: Genomic DNA  
US-08-936-165A-92

Query Match 67.8%; Score 15.6; DB 4; Length 1142;  
Best Local Similarity 81.8%; Pred. No. 96;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23  
DB 797 GTACCATGTGTCAAAATTGAC 818

## RESULT 7

US-09-657-453A-17  
/ Sequence 17, Application US/09657453A  
/ Patent No. 6458591  
/ GENERAL INFORMATION:  
/ APPLICANT: Brett P. Monia  
/ APPLICANT: Jacqueline Wyatt  
/ TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHORYLASE KINASE ALPHA 2 EXPRESS  
/ FILE REFERENCE: RTS-0136  
/ CURRENT APPLICATION NUMBER: US/09/657,453A  
/ CURRENT FILING DATE: 2000-09-07  
/ NUMBER OF SEQ ID NOS: 105  
/ SEQ ID NO 17  
/ LENGTH: 4392  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (172)...(3879)  
US-09-657-453A-17

Query Match 67.8%; Score 15.6; DB 4; Length 4392;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23

DB 3702 GGACCGGTGTCAATTGGTGCC 3723  
RESULT 8  
US-09-657-453A-3  
/ Sequence 3, Application US/09657453A  
/ Patent No. 6458591  
/ GENERAL INFORMATION:  
/ APPLICANT: Brett P. Monia  
/ APPLICANT: Jacqueline Wyatt  
/ TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHORYLASE KINASE ALPHA 2 EXPRESS  
/ FILE REFERENCE: RTS-0136  
/ CURRENT APPLICATION NUMBER: US/09/657,453A  
/ CURRENT FILING DATE: 2000-09-07  
/ NUMBER OF SEQ ID NOS: 105  
/ SEQ ID NO 3  
/ LENGTH: 4566  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (127)...(3834)  
US-09-657-453A-3

Query Match 67.8%; Score 15.6; DB 4; Length 4566;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23  
DB 3657 GGACCGGTGTCAATTGGTGCC 3678

## RESULT 9

US-09-103-840A-2/c  
/ Sequence 2, Application US/09103840A  
/ Patent No. 6294328  
/ GENERAL INFORMATION:  
/ APPLICANT: FLEISCHMAN, Robert D.  
/ APPLICANT: WHITE, Owen R.  
/ APPLICANT: FRASER, Claire M.  
/ APPLICANT: VENTER, John C.  
/ TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
/ FILE REFERENCE: 24366-20007.00  
/ CURRENT APPLICATION NUMBER: US/09/103,840A  
/ CURRENT FILING DATE: 1998-06-24  
/ NUMBER OF SEQ ID NOS: 2  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 2  
/ LENGTH: 4403765  
/ TYPE: DNA  
/ ORGANISM: Mycobacterium tuberculosis  
/ FEATURE:  
/ OTHER INFORMATION: CDC 1551  
/ OTHER INFORMATION: "n" bases at various positions throughout the sequence  
/ OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 67.8%; Score 15.6; DB 3; Length 4403765;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23  
DB 2679869 GGATCATGTGTCCGCTATGCC 2679848

## RESULT 10

US-09-103-840A-1/c  
/ Sequence 1, Application US/09103840A  
/ Patent No. 6294328

```
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          67.8%; Score 15.6; DB 3; Length 4411529;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATGCC 23
Db 2682574 GGATCATGTGTCCGCGTATGCC 2682553

RESULT 11
US-09-390-234-13
; Sequence 13, Application US/09390234
; Patent No. 6365390
; GENERAL INFORMATION:
; APPLICANT: Blum, David L.
; APPLICANT: Katsava, Irina
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; TITLE OF INVENTION: Phenolic Acid Esterases, Coding Sequences and Methods
; FILE REFERENCE: 67-98
; CURRENT APPLICATION NUMBER: US/09/390,234
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: US 60/099,136
; EARLIER FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2722
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2611)
US-09-390-234-13

Query Match          67.0%; Score 15.4; DB 4; Length 2722;
Best Local Similarity 94.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACT 17
Db 1687 CGGACATGTGTCAACT 1703

RESULT 12
US-09-603-311-13
; Sequence 13, Application US/09603311
; Patent No. 6602700
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Azain, Michael J.
; APPLICANT: Davies, Edward T.
; APPLICANT: Shah, Ashit K.
```

```
; APPLICANT: Blum, David L.
; APPLICANT: Katsava, Irina
; TITLE OF INVENTION: Phenolic Acid Esterases, Coding Sequences and Methods
; FILE REFERENCE: 67-98A
; CURRENT APPLICATION NUMBER: US/09/603,311
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/099,136
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 09/390,324
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2722
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2611)
US-09-603-311-13

Query Match          67.0%; Score 15.4; DB 4; Length 2722;
Best Local Similarity 94.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACT 17
Db 1687 CGGACATGTGTCAACT 1703

RESULT 13
US-09-442-021-5/c
; Sequence 5, Application US/09442021
; Patent No. 5632980
; GENERAL INFORMATION:
; APPLICANT: Yadav, Narendra S.
; TITLE OF INVENTION: Binary Viral Expression System in Plants
; FILE REFERENCE: CU1127 US CIP
; CURRENT APPLICATION NUMBER: US/09/442,021
; CURRENT FILING DATE: 1999-11-17
; EARLIER APPLICATION NUMBER: 60/063,504
; EARLIER FILING DATE: 1997-October-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-442-021-5

Query Match          66.1%; Score 15.2; DB 4; Length 33;
Best Local Similarity 85.0%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ACCATGTGTCAACTTATGCC 23
Db 22 ACCATGTGTCCGCGCATGCC 3

RESULT 14
US-09-621-976-1327
; Sequence 1327, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1327
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 208..480
US-09-621-976-1327
```

```
Query Match 66.1%; Score 15.2; DB 4; Length 480;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 GACCATGTGTCAACTTATGC 22
|||||
Db 235 GACCATGTGACATCTGATGC 254
```

```
RESULT 15
US-09-702-705-1094/c
; Sequence 1094, Application US/09702705
; Patent No. 650401C
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1094
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(610)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-705-1094
```

```
Query Match 66.1%; Score 15.2; DB 4; Length 610;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTAT 20
|||||
Db 223 CTGACCATGTTTCAACTTCT 204
```

```
RESULT 16
US-09-736-457-1094/c
; Sequence 1094, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
```

```
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1094
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(610)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-1094
```

```
Query Match 66.1%; Score 15.2; DB 4; Length 610;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTAT 20
|||||
Db 223 CTGACCATGTTTCAACTTCT 204
```

```
RESULT 17
US-09-614-124B-1094/c
; Sequence 1094, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1094
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(610)
; OTHER INFORMATION: n = A,T,C or G
US-09-614-124B-1094
```

```
Query Match 66.1%; Score 15.2; DB 4; Length 610;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTAT 20
|||||
Db 223 CTGACCATGTTTCAACTTCT 204
```

```
RESULT 18
US-09-671-325-1094/c
; Sequence 1094, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
```



C 88	15.6	67.8	566	10	US-09-986-480-121	Sequence 121, App	161	15.2	66.1	1048	13	US-10-027-632-122252	Sequence 122252,
C 89	15.6	67.8	685	13	US-10-027-632-251723	Sequence 251723, A	162	15.2	66.1	1048	16	US-10-027-632-122252	Sequence 122252,
C 90	15.6	67.8	685	16	US-10-027-632-251723	Sequence 251723, A	163	15.2	66.1	1120	10	US-09-890-688-155	Sequence 155, App
C 91	15.6	67.8	693	13	US-10-027-632-259559	Sequence 259559, A	c 164	15.2	66.1	1134	13	US-10-282-122A-28069	Sequence 28069, A
C 92	15.6	67.8	693	16	US-10-027-632-259559	Sequence 259559, A	c 165	15.2	66.1	1138	13	US-10-425-114-9614	Sequence 9614, App
C 93	15.6	67.8	1142	9	US-09-939-980-92	Sequence 92, Appl	c 166	15.2	66.1	1142	13	US-10-424-599-8364	Sequence 8364, App
C 94	15.6	67.8	1561	13	US-10-424-599-82780	Sequence 82780, A	c 167	15.2	66.1	1539	15	US-10-106-698-548	Sequence 548, App
C 95	15.6	67.8	1619	15	US-10-174-363-25	Sequence 25, Appl	c 168	15.2	66.1	1554	13	US-10-425-114-30158	Sequence 30158, A
C 96	15.6	67.8	1646	13	US-10-282-122A-29374	Sequence 29374, A	c 169	15.2	66.1	1570	13	US-10-027-632-252127	Sequence 252127, A
C 97	15.6	67.8	2505	13	US-10-027-632-101493	Sequence 101493, A	c 170	15.2	66.1	1570	13	US-10-027-632-252128	Sequence 252128, A
C 98	15.6	67.8	2505	16	US-10-027-632-101493	Sequence 101493, A	c 171	15.2	66.1	1570	13	US-10-027-632-252129	Sequence 252129, A
C 99	15.6	67.8	5056	10	US-09-919-039-230	Sequence 230, App	c 172	15.2	66.1	1570	16	US-10-027-632-252127	Sequence 252127, A
C 100	15.6	67.8	5423	15	US-10-252-157-11	Sequence 11, Appl	c 173	15.2	66.1	1570	16	US-10-027-632-252128	Sequence 252128, A
C 101	15.6	67.8	8957	10	US-09-764-891-9239	Sequence 9239, App	c 174	15.2	66.1	1570	16	US-10-027-632-252127	Sequence 252127, A
C 102	15.6	67.8	8957	15	US-10-091-572-815	Sequence 815, App	c 175	15.2	66.1	1620	9	US-09-938-842A-960	Sequence 960, App
C 103	15.6	67.8	164875	16	US-10-085-117-322	Sequence 322, App	c 176	15.2	66.1	1620	11	US-09-938-842A-960	Sequence 960, App
C 104	15.6	67.0	442	9	US-09-770-447-873	Sequence 873, App	c 177	15.2	66.1	1646	16	US-10-087-080-25	Sequence 25, Appl
C 105	15.4	67.0	663	13	US-10-027-632-227010	Sequence 227010, A	c 178	15.2	66.1	1743	13	US-10-027-632-266098	Sequence 266098, A
C 106	15.4	67.0	663	13	US-10-027-632-227011	Sequence 227011, A	c 179	15.2	66.1	1743	13	US-10-027-632-266099	Sequence 266099, A
C 107	15.4	67.0	663	16	US-10-027-632-227011	Sequence 227011, A	c 180	15.2	66.1	1743	13	US-10-027-632-266100	Sequence 266100, A
C 108	15.4	67.0	663	16	US-10-027-632-227011	Sequence 227011, A							
C 109	15.4	67.0	688	13	US-10-027-632-206967	Sequence 206967, A							
C 110	15.4	67.0	688	16	US-10-027-632-206967	Sequence 206967, A							
C 111	15.4	67.0	1191	13	US-10-027-632-123199	Sequence 123199, A							
C 112	15.4	67.0	1191	16	US-10-027-632-123199	Sequence 123199, A							
C 113	15.4	67.0	1319	13	US-10-027-632-209869	Sequence 209869, A							
C 114	15.4	67.0	1319	16	US-10-027-632-209869	Sequence 209869, A							
C 115	15.4	67.0	7928	16	US-09-764-847-1535	Sequence 1535, App							
C 116	15.4	67.0	7928	9	US-10-092-154-1535	Sequence 1535, App							
C 117	15.2	66.1	113	9	US-09-796-692-9129	Sequence 9129, App							
C 118	15.2	66.1	113	9	US-09-796-692-9129	Sequence 9129, App							
C 119	15.2	66.1	113	15	US-10-040-862-9138	Sequence 9138, App							
C 120	15.2	66.1	113	15	US-10-040-862-9138	Sequence 9138, App							
C 121	15.2	66.1	113	15	US-10-040-862-9138	Sequence 9138, App							
C 122	15.2	66.1	113	16	US-10-057-475B-9129	Sequence 9129, App							
C 123	15.2	66.1	113	16	US-10-057-475B-9138	Sequence 9138, App							
C 124	15.2	66.1	113	16	US-10-154-884B-9129	Sequence 9129, App							
C 125	15.2	66.1	201	16	US-10-154-884B-9138	Sequence 9138, App							
C 126	15.2	66.1	201	9	US-09-864-761-18750	Sequence 18750, A							
C 127	15.2	66.1	201	9	US-09-864-761-21744	Sequence 21744, A							
C 128	15.2	66.1	307	9	US-09-783-590-1357	Sequence 1357, App							
C 129	15.2	66.1	385	10	US-09-918-995-16455	Sequence 16455, A							
C 130	15.2	66.1	392	13	US-10-027-632-287042	Sequence 287042, A							
C 131	15.2	66.1	392	13	US-10-027-632-287042	Sequence 287042, A							
C 132	15.2	66.1	392	16	US-10-027-632-287042	Sequence 287042, A							
C 133	15.2	66.1	392	16	US-10-027-632-287043	Sequence 287043, A							
C 134	15.2	66.1	417	9	US-09-983-965-1806	Sequence 1806, App							
C 135	15.2	66.1	453	10	US-09-918-995-29361	Sequence 29361, A							
C 136	15.2	66.1	483	10	US-09-918-995-13935	Sequence 13935, A							
C 137	15.2	66.1	547	13	US-10-027-632-304158	Sequence 304158, A							
C 138	15.2	66.1	547	16	US-10-027-632-304158	Sequence 304158, A							
C 139	15.2	66.1	558	13	US-10-027-632-279476	Sequence 279476, A							
C 140	15.2	66.1	558	16	US-10-027-632-279476	Sequence 279476, A							
C 141	15.2	66.1	603	13	US-10-424-599-92417	Sequence 92417, A							
C 142	15.2	66.1	610	9	US-09-736-457-1094	Sequence 1094, App							
C 143	15.2	66.1	610	9	US-09-902-941-1094	Sequence 1094, App							
C 144	15.2	66.1	610	9	US-09-849-626-1094	Sequence 1094, App							
C 145	15.2	66.1	610	13	US-10-283-017-1094	Sequence 1094, App							
C 146	15.2	66.1	610	15	US-10-017-754-1094	Sequence 1094, App							
C 147	15.2	66.1	610	15	US-10-113-872-1094	Sequence 1094, App							
C 148	15.2	66.1	611	13	US-10-027-632-214732	Sequence 214732, A							
C 149	15.2	66.1	611	16	US-10-027-632-214733	Sequence 214733, A							
C 150	15.2	66.1	611	16	US-10-027-632-214732	Sequence 214732, A							
C 151	15.2	66.1	621	13	US-10-027-632-214733	Sequence 214733, A							
C 152	15.2	66.1	621	16	US-10-027-632-259824	Sequence 259824, A							
C 153	15.2	66.1	621	16	US-10-027-632-259824	Sequence 259824, A							
C 154	15.2	66.1	663	9	US-09-974-300-6876	Sequence 5876, App							
C 155	15.2	66.1	691	9	US-09-974-300-821	Sequence 821, App							
C 156	15.2	66.1	735	13	US-10-027-632-103753	Sequence 103753, A							
C 157	15.2	66.1	735	16	US-10-027-632-103753	Sequence 103753, A							
C 158	15.2	66.1	746	13	US-10-027-632-12032	Sequence 12032, A							
C 159	15.2	66.1	746	16	US-10-027-632-12032	Sequence 12032, A							
C 160	15.2	66.1	927	16	US-10-369-493-25312	Sequence 25312, A							
			994	9	US-09-925-301-231	Sequence 231, App							

ALIGNMENTS

RESULT 1

US-09-877-819B-34  
; Sequence 34, Application US/09877819B  
; Publication No. US20030190609A1  
; GENERAL INFORMATION:  
; APPLICANT: White, Scott  
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing  
; FILE REFERENCE: S-94,664  
; CURRENT APPLICATION NUMBER: US/09/877,819B  
; CURRENT FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 23  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Primer sequence  
US-09-877-819B-34

Query Match 100.0%; Score 23; DB 10; Length 23;  
Best Local Similarity 100.0%; Pred. NO. 0.13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGGACCATGTGTCACCTATGCC 23

Db 1 CGGACCATGTGTCACCTATGCC 23

RESULT 2

US-10-297-068-443  
; Sequence 443, Application US/10297068  
; Publication No. US20030228585A1  
; GENERAL INFORMATION:  
; APPLICANT: INOKO, Hidetoshi  
; APPLICANT: ICHIHARA, Taeko  
; APPLICANT: KAGIYA, Taeko  
; APPLICANT: Matsumura, Yoshiyuki  
; APPLICANT: MORIYA, Shogo  
; APPLICANT: NISHIDA, Michio  
; TITLE OF INVENTION: KIT AND METHOD FOR DETERMINING HLA TYPES  
; FILE REFERENCE: 1314024174  
; CURRENT APPLICATION NUMBER: US/10/297,068  
; CURRENT FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: JP 2000-164798  
; PRIOR FILING DATE: 2000-06-01  
; NUMBER OF SEQ ID NOS: 1298

```
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 443
/ LENGTH: 26
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:primer
US-10-297-068-443

Query Match          100.0%; Score 23; DB 16; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
Db 4 CGGACCATGTGTCAACTTATGCC 26

RESULT 3
US-09-877-819B-55
/ Sequence 55, Application US/09877819B
/ Publication No. US20030190609A1
/ GENERAL INFORMATION:
/ APPLICANT: Torney, David
/ TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
/ FILE REFERENCE: S-94,664
/ CURRENT APPLICATION NUMBER: US/09/877,819B
/ CURRENT FILING DATE: 2001-06-07
/ NUMBER OF SEQ ID NOS: 55
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 55
/ LENGTH: 254
/ TYPE: DNA
/ ORGANISM: Human HLA
US-09-877-819B-55

Query Match          100.0%; Score 23; DB 10; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
Db 8 CGGACCATGTGTCAACTTATGCC 30

RESULT 4
US-09-796-692-143
/ Sequence 143, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
/ CURRENT FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 328
/ LENGTH: 267
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-328

Query Match          100.0%; Score 23; DB 9; Length 267;
```

```
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 143
/ LENGTH: 267
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-143

Query Match          100.0%; Score 23; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
Db 129 CGGACCATGTGTCAACTTATGCC 151

RESULT 5
US-09-796-692-328
/ Sequence 328, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
/ CURRENT FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 328
/ LENGTH: 267
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-328

Query Match          100.0%; Score 23; DB 9; Length 267;
```

```
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 129 CGGACCATGTGTCAACTTATGCC 151

RESULT 6
US-09-796-692-7075
; Sequence 7075, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7075

Query Match 100.0%; Score 23; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 129 CGGACCATGTGTCAACTTATGCC 151

RESULT 7
US-10-040-862-143
; Sequence 143, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
```

```
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-0135200S
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-143

Query Match 100.0%; Score 23; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 129 CGGACCATGTGTCAACTTATGCC 151

RESULT 8
US-10-040-862-328
; Sequence 328, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-0135200S
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-04-28
```

```
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,064
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-328
```

```
Query Match 100.0%; Score 23; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 129 CGGACCATGTGTCAACTTATGCC 151
```

```
RESULT 9
US-10-040-862-7075
; Sequence 7075, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-143
```

```
Query Match 100.0%; Score 23; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 129 CGGACCATGTGTCAACTTATGCC 151
```

```
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7075
```

```
Query Match 100.0%; Score 23; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 129 CGGACCATGTGTCAACTTATGCC 151
```

```
RESULT 10
US-10-057-475B-143
; Sequence 143, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-143
```

```
Query Match 100.0%; Score 23; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7075
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-7075

Query Match 100.0%; Score 23; DB 16; Length 267;
Best Local Similarity 100.0%; Pred.No.0.18; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 129 CGGACCATGTGTCAACTTATGCC 151

RESULT 13
US-10-154-884B-143
; Sequence 143, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22

```

; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 143  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-143

Query Match 100.0%; Score 23; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
DB 129 CGGACCATGTGTCAACTTATGCC 151

## RESULT 14

US-10-154-884B-328  
; Sequence 328, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

; FILE REFERENCE: 014058-013521US

; CURRENT APPLICATION NUMBER: US/10/154,884B

; CURRENT FILING DATE: 2002-05-23

; PRIOR APPLICATION NUMBER: US 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: US 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: US 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: US 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: US 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: US 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/222,903

; PRIOR FILING DATE: 2000-08-03

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 11290

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 328

; LENGTH: 267

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-154-884B-328

Query Match 100.0%; Score 23; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
|||||  
DB 129 CGGACCATGTGTCAACTTATGCC 151

```
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 42
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-42
```

```
Query Match 100.0%; Score 23; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY - CGGACCATGTGTCAACTTATGCC 23
      |||||
DB 134 CGGACCATGTGTCAACTTATGCC 156
```

```
RESULT 17
US-09-796-692-3554
/ Sequence 3554, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ FILE REFERENCE: 2077,001200
/ CURRENT APPLICATION NUMBER: US/09796,692
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 42
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-42
```

```
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3554
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-3554
```

```
Query Match 100.0%; Score 23; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
      |||||
DB 134 CGGACCATGTGTCAACTTATGCC 156
```

```
RESULT 18
US-10-040-862-42
/ Sequence 42, Application US/10040862
/ Publication No. US20030078396A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 42
/ LENGTH: 272
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-040-862-42
```

```
Query Match 100.0%; Score 23; DB 15; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
      |||||
DB 134 CGGACCATGTGTCAACTTATGCC 156
```

```
RESULT 19
US-10-040-862-3554
; Sequence 3554, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-3554

Query Match 100.0%; Score 23; DB 15; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 134 CGGACCATGTGTCAACTTATGCC 156

RESULT 20
US-10-057-475B-42
; Sequence 42, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; APPLICANT: McNeill, Patricia Dianne

Query Match 100.0%; Score 23; DB 15; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 134 CGGACCATGTGTCAACTTATGCC 156

RESULT 21
US-10-057-475B-3554
; Sequence 3554, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; APPLICANT: McNeill, Patricia Dianne
```

Query Match 100.0%; Score 23; DB 16; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
Db 134 CGGACCATGTGTCAACTTATGCC 156  
RESULT 22  
US-10-154-884B-42  
Sequence 42, Application US/10154884B  
Publication No. US20040005561A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc W.  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
FILE REFERENCE: 014058-013521US  
CURRENT APPLICATION NUMBER: US/10/154,884B  
CURRENT FILING DATE: 2002-05-23  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11290  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3554  
LENGTH: 272  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-057-475B-3554

Query Match 100.0%; Score 23; DB 16; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
Db 134 CGGACCATGTGTCAACTTATGCC 156

RESULT 22  
US-10-154-884B-42  
Sequence 42, Application US/10154884B  
Publication No. US20040005561A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc W.  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
FILE REFERENCE: 014058-013521US  
CURRENT APPLICATION NUMBER: US/10/154,884B  
CURRENT FILING DATE: 2002-05-23  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11290  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 42  
LENGTH: 272  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-154-884B-42

Query Match 100.0%; Score 23; DB 16; Length 272;

Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCATGTGTCAACTTATGCC 23  
Db 134 CGGACCATGTGTCAACTTATGCC 156

RESULT 23  
US-10-154-884B-3554  
Sequence 3554, Application US/10154884B  
Publication No. US20040005561A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc W.  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
FILE REFERENCE: 014058-013521US  
CURRENT APPLICATION NUMBER: US/10/154,884B  
CURRENT FILING DATE: 2002-05-23  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11290  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3554  
LENGTH: 272  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-154-884B-3554

Query Match 100.0%; Score 23; DB 16; Length 272;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23  
Db 134 CGGACCATGTGTCAACTTATGCC 156

RESULT 24  
US-09-796-692-280/c  
Sequence 280, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
APPLICANT: Algate, Paul A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692

```
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-03-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-280
```

```
Query Match 100.0%; Score 23; DB 9; Length 294;
Best Local Similarity 100.0%; Pred.No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 139 CGGACCATGTGTCAACTTATGCC 117
```

```
RESULT 25
US-09-796-692-530/c
; Sequence 530, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2003-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
```

```
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-530
```

```
Query Match 100.0%; Score 23; DB 9; Length 294;
Best Local Similarity 100.0%; Pred.No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
DB 139 CGGACCATGTGTCAACTTATGCC 117
```

```
RESULT 26
US-09-796-692-3270/c
; Sequence 3270, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-3270
```

```
Query Match 100.0%; Score 23; DB 9; Length 294;
Best Local Similarity 100.0%; Pred.No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
```



```
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-3270

Query Match      100.0%; Score 23; DB 15; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117

RESULT 30
US-10-057-475B-280/c
; Sequence 280, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US 60/186,126
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/200,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-3270

Query Match      100.0%; Score 23; DB 15; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117

RESULT 30
US-10-057-475B-280/c
; Sequence 280, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-280

Query Match      100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117

RESULT 31
US-10-057-475B-530/c
; Sequence 530, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-530

Query Match      100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 139 CGGACCATGTGTCAACTTATGCC 117
```



## RESULT 32

```
US-10-057-475B-3270/c
; Sequence 3270, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10975
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-3270
```

```
Query Match          100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
Db 139 CGGACCATGTGTCAACTTATGCC 117
```

## RESULT 33

```
US-10-154-884B-280/c
; Sequence 280, Application US/10154884B
; Publication No. US200400005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
```

```
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-280
```

```
Query Match          100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
Db 139 CGGACCATGTGTCAACTTATGCC 117
```

## RESULT 34

```
US-10-154-884B-530/c
; Sequence 530, Application US/10154884B
; Publication No. US200400005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-530
```

```
Query Match      100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 139 CGGACCATGTGTCAACTTATGCC 117
```

## RESULT 35

```
US-10-154-884B-3270/C
; Sequence 3270, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3270
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-3270
```

```
Query Match      100.0%; Score 23; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 139 CGGACCATGTGTCAACTTATGCC 117
```

## RESULT 36

```
US-09-918-995-36460
; Sequence 36460, Application US/09918995
```

```
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36460
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-36460
```

```
Query Match      100.0%; Score 23; DB 10; Length 410;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 152 CGGACCATGTGTCAACTTATGCC 174
```

## RESULT 37

```
US-09-918-995-37140
; Sequence 37140, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37140
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-37140
```

```
Query Match      100.0%; Score 23; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGACCATGTGTCAACTTATGCC 23
    |||||
DB 225 CGGACCATGTGTCAACTTATGCC 247
```

## RESULT 38

```
US-10-102-524-715
; Sequence 715, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; THERAPY AND DIAGNOSIS OF KIDNEY CANCER
```

```
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 715
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-715

Query Match      100.0%; Score 23; DB 15; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 140 CGGACCATGTGTCAACTTATGCC 162

RESULT 39
US-10-084-817-88
; Sequence 88, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 88
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1090035.1
US-10-084-817-88

Query Match      100.0%; Score 23; DB 15; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 173 CGGACCATGTGTCAACTTATGCC 195

RESULT 40
US-10-220-120-138
; Sequence 138, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.

; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: KONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YI, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINGCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-16;
; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17; 2000-05-17;
; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 138
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:1169865.1:2000MAY01
US-10-220-120-138

Query Match      100.0%; Score 23; DB 13; Length 1259;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 315 CGGACCATGTGTCAACTTATGCC 337

RESULT 41
US-09-925-302-351
; Sequence 351, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
```

```
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1294)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1318)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1329)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-302-351
```

```
Query Match 100.0%; Score 23; DB 9; Length 1348;
Best Local Similarity 100.0%; Pred. No. 0.22; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 23; Conservative 0;
QY 1 CGGACCATGTGTCACACTTATGCC 23
Db 253 CGGACCATGTGTCACACTTATGCC 275
```

```
RESULT 42
US-09-925-302-351
; Sequence 351, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1294)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1318)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1329)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-302-351
```

```
Query Match 100.0%; Score 23; DB 13; Length 1348;
Best Local Similarity 100.0%; Pred. No. 0.22; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 23; Conservative 0;
QY 1 CGGACCATGTGTCACACTTATGCC 23
Db 253 CGGACCATGTGTCACACTTATGCC 275
```

```
RESULT 43
US-09-960-706-1043
; Sequence 1043, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
; US-09-960-706-1043
```

```
Query Match 100.0%; Score 23; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.3; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 23; Conservative 0;
QY 1 CGGACCATGTGTCACACTTATGCC 23
Db 6334 CGGACCATGTGTCACACTTATGCC 6356
```

```
RESULT 44
US-09-873-319-691
; Sequence 691, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
```

```
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 691
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-873-319-691

Query Match      100.0%; Score 23; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGGACCATGTGTCAACTTATGCC 23
Db      5334 CGGACCATGTGTCAACTTATGCC 6356

RESULT 45
US-09-258-133-13
; Sequence 13, Application US/09258133
; Publication No. US20030044779A1
; GENERAL INFORMATION:
; APPLICANT: Goelet, Philip
; APPLICANT: Knapp, Michael
; APPLICANT: Anderson, Stephen
; TITLE OF INVENTION: REAGENT FOR NUCLEIC ACID TYPING BY POLYMERASE EXTENSION
; FILE REFERENCE: 13017-2
; CURRENT APPLICATION NUMBER: US/09/258,133
; CURRENT FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 07/775,786
; PRIOR FILING DATE: 1991-10-11
; PRIOR APPLICATION NUMBER: 07/664,837
; PRIOR FILING DATE: 1991-03-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic primer for human HLA DPAl
US-09-258-133-13

Query Match      87.0%; Score 20; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGGACCATGTGTCAACTTAT 20
Db      2 CGGACCATGTGTCAACTTAT 21

RESULT 46
US-10-425-114-2710
; Sequence 2710, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33564
; LENGTH: 2212

US-10-425-114-33564/c
; Sequence 33564, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33564
; LENGTH: 2212

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2710
; LENGTH: 2351
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 7002221995_FLI
US-10-425-114-2710

Query Match      78.3%; Score 18; DB 13; Length 2351;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGGACCATGTGTCAACTT 18
Db      350 CGGACCATGTGTCAACTT 367

RESULT 47
US-10-425-114-16978/c
; Sequence 16978, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16978
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3068-049-C2_FLI
US-10-425-114-16978

Query Match      77.4%; Score 17.8; DB 13; Length 1438;
Best Local Similarity 90.5%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGACCATGTGTCAACTTATGC 22
Db      914 GGACCATGTGTCAACTTATGC 894

RESULT 48
US-10-425-114-33564/c
; Sequence 33564, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33564
; LENGTH: 2212
```

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017134B09_FLI
US-10-425-114-33564

Query Match          77.4%; Score 17.8; DB 13; Length 2212;
Best Local Similarity 90.5%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 22
    |||||
Db 1259 GGACCATGTGTCAACTTATGCC 1239

RESULT 49
US-10-027-632-269067
; Sequence 269067, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269067
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269067

Query Match          74.8%; Score 17.2; DB 13; Length 549;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 23
    |||||
Db 418 GGAATATGTGTAACTTATGCC 439

RESULT 50
US-10-027-632-269067
; Sequence 269067, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269067
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269067

Query Match          74.8%; Score 17.2; DB 13; Length 549;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 23
    |||||
Db 418 GGAATATGTGTAACTTATGCC 439

RESULT 51
US-10-027-632-269066
; Sequence 269066, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269066
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269066

Query Match          74.8%; Score 17.2; DB 13; Length 640;
Best Local Similarity 86.4%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 23
    |||||
Db 509 GGAATATGTGTAACTTATGCC 530

RESULT 52
US-10-027-632-269066
; Sequence 269066, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-10-027-632-269067

Query Match          74.8%; Score 17.2; DB 16; Length 549;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 23
    |||||
Db 418 GGAATATGTGTAACTTATGCC 439

RESULT 51
US-10-027-632-269066
; Sequence 269066, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269066
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269066

Query Match          74.8%; Score 17.2; DB 13; Length 640;
Best Local Similarity 86.4%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGACCATGTGTCAACTTATGCC 23
    |||||
Db 509 GGAATATGTGTAACTTATGCC 530

RESULT 52
US-10-027-632-269066
; Sequence 269066, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269066
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269066
```

```
Query Match 74.8%; Score 17.2; DB 16; Length 640;
Best Local Similarity 86.4%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 GGACCATGTGTCAACTTATGCC 23
||| ||||| ||||| |||||
Db 509 GGAATATGTGTTAACTTATGCC 530
```

```
RESULT 53
US-10-027-632-269068
; Sequence 269068, Application US/10027632
; Publication No. US2002019837LAI
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269068
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269068
```

```
Query Match 74.8%; Score 17.2; DB 13; Length 1280;
Best Local Similarity 86.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 GGACCATGTGTCAACTTATGCC 23
||| ||||| ||||| |||||
Db 509 GGAATATGTGTTAACTTATGCC 530
```

```
RESULT 54
US-10-027-632-269068
; Sequence 269068, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269068
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269068
```

```
Query Match 74.8%; Score 17.2; DB 16; Length 1280;
Best Local Similarity 86.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 GGACCATGTGTCAACTTATGCC 23
||| ||||| ||||| |||||
Db 509 GGAATATGTGTTAACTTATGCC 530
```

```
RESULT 55
US-09-887-576-839/c
; Sequence 839, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 839
```

```

; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-839

Query Match          73.0%; Score 16.8; DB 9; Length 2000;
Best Local Similarity 90.0%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGACCATGTGTCAACTTATG 21
Db 820 GGACCATGTTCACCTATG 801

RESULT 56
US-10-425-114-32493
; Sequence 32493, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32493
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493

Query Match          72.2%; Score 16.6; DB 13; Length 2445;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 836 CGAACATTTGTCAAATTATGCC 858

RESULT 57
US-10-425-114-31957
; Sequence 31957, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31957
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73238H12_FLI
US-10-425-114-31957
```

```

Query Match          72.2%; Score 16.6; DB 13; Length 2537;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 1009 CGAACATTTGTCAAATTATGCC 1031

RESULT 58
US-10-425-114-31061
; Sequence 31061, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31061
; LENGTH: 3607
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73122F03_FLI
US-10-425-114-31061

Query Match          72.2%; Score 16.6; DB 13; Length 3607;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 2081 CGAACATTTGTCAAATTATGCC 2103

RESULT 59
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match          72.2%; Score 16.6; DB 15; Length 3673778;
Best Local Similarity 82.6%; Pred. No. 6.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23
Db 3298061 CGAACATATATCAACTTATACC 3298039
```



```
RESULT 60
US-09-294-093B-915
; Sequence 915, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 915
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700343508H1
US-09-294-093B-915
```

```
Query Match          70.4%; Score 16.2; DB 9; Length 286;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 CGGACCATCTGTCAACTTATG 21
         |||||
Db       91 CGGACCATCTGTGAACCTTAAG 111
```

```
Search completed: April 20, 2004, 12:55:04
Job time : 148.112 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1463.75 Seconds  
(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819B-34

Perfect score: 23

Sequence: 1 cggaccatgtgcacttatgcc 23

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_eston:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rnd:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	159	14	T91602
2	23	100.0	168	14	CD686854
3	23	100.0	177	10	BG025808
4	23	100.0	229	10	B3140710
C					

5	23	100.0	256	10	BE159476
6	23	100.0	265	10	BE242285
7	23	100.0	289	9	AA360953
8	23	100.0	290	14	CD706205
9	23	100.0	295	14	CD694567
10	23	100.0	318	10	AW404262
11	23	100.0	334	14	CD704658
12	23	100.0	335	12	BM772902
13	23	100.0	338	9	AA323639
14	23	100.0	365	14	T94759
15	23	100.0	387	14	CB267061
16	23	100.0	390	14	CD709023
17	23	100.0	405	9	AA244273
18	23	100.0	408	12	BM834672
19	23	100.0	415	9	AA838010
20	23	100.0	418	10	BF378520
21	23	100.0	422	13	C03540
22	23	100.0	423	14	CD102141
23	23	100.0	427	9	AW085969
24	23	100.0	437	12	BM767805
25	23	100.0	439	10	AW406315
26	23	100.0	443	14	CD687150
27	23	100.0	444	14	CD704753
28	23	100.0	464	14	CD705070
29	23	100.0	469	12	BM766631
30	23	100.0	483	14	CD692032
31	23	100.0	484	10	BF819626
32	23	100.0	491	12	BM694247
33	23	100.0	494	14	CD698716
34	23	100.0	511	14	CD707609
35	23	100.0	514	14	CD698819
36	23	100.0	515	12	BM769742
37	23	100.0	519	12	BG541135
38	23	100.0	526	14	CD705087
39	23	100.0	527	14	CD684349
40	23	100.0	530	14	CD708787
41	23	100.0	533	14	CD687511
42	23	100.0	534	12	BG756165
43	23	100.0	540	14	CD695435
44	23	100.0	544	10	BE874055
45	23	100.0	545	13	BQ267707
46	23	100.0	545	13	BQ267707
47	23	100.0	546	10	AW351777
48	23	100.0	547	14	CD706950
49	23	100.0	547	10	AW406086
50	23	100.0	548	14	CD696409
51	23	100.0	557	12	BG535978
52	23	100.0	557	14	CD687141
53	23	100.0	559	14	CD700948
54	23	100.0	563	14	CB265399
55	23	100.0	571	14	CA942442
56	23	100.0	579	12	BM831052
57	23	100.0	582	12	BM737984
58	23	100.0	585	9	AV706521
59	23	100.0	585	13	BQ783392
60	23	100.0	588	12	B1911442
61	23	100.0	588	14	CD698528
62	23	100.0	590	14	CD707303
63	23	100.0	592	12	BM876262
64	23	100.0	594	12	BG926106
65	23	100.0	595	14	CD693703
66	23	100.0	597	14	CD684394
67	23	100.0	598	14	CD691317
68	23	100.0	598	14	CD695219
69	23	100.0	599	13	BQ270752
70	23	100.0	601	10	B725429
71	23	100.0	602	13	BQ783579
72	23	100.0	602	14	CD699993
73	23	100.0	605	12	BM820452
74	23	100.0	608	14	CD706661
75	23	100.0	610	14	CA405960
76	23	100.0	618	14	CD686880
77	23	100.0	623	10	BF128930

```

78 23 100.0 626 12 BG545380 602572611
79 23 100.0 627 12 BG715375 602675513
80 23 100.0 627 14 CB267128 1006034 K
81 23 100.0 628 14 CD702973 EST19580
82 23 100.0 629 14 CD691540 EST8063 h
83 23 100.0 635 10 BE712825 M80-HT069
84 23 100.0 635 12 BG709234 602674620
85 23 100.0 643 12 SM767502 K-EST0049
86 23 100.0 645 12 BI769913 603060181
87 23 100.0 646 14 CD689742 EST6265 h
88 23 100.0 658 9 AV763677 AV763677
89 23 100.0 659 14 CD687077 EST3598 h
90 23 100.0 663 14 CD701461 EST17885
91 23 100.0 667 9 AL048565 DRF2p586P
92 23 100.0 669 13 BU071266 IM41B05.X
93 23 100.0 671 12 BG540027 602568650
94 23 100.0 672 14 CD692599 EST9122 h
95 23 100.0 676 12 BG570300 602590723
96 23 100.0 683 9 AV759427 AV759427
97 23 100.0 684 14 CD693272 EST9795 h
98 23 100.0 685 10 BE973719 601680707
99 23 100.0 693 12 BG685335 602637767
100 23 100.0 697 14 CD708187 EST24714
101 23 100.0 698 12 BI834690 603090472
102 23 100.0 700 14 CD696900 EST13423
103 23 100.0 702 12 BG541647 602571219
104 23 100.0 704 9 AV707666 AV707666
105 23 100.0 706 12 BG431973 602498115
106 23 100.0 708 12 BG539584 602567527
107 23 100.0 708 14 N25099 YX1812.1 r1
108 23 100.0 709 12 BG758164 602712334
109 23 100.0 720 10 BF976100 602244936
110 23 100.0 724 10 BE435689 HTM1-360P
111 23 100.0 732 12 BG699921 602681285
112 23 100.0 732 14 CB997879 AGENCOURT
113 23 100.0 734 9 AV733676 AV733676
114 23 100.0 735 12 BG686058 602638627
115 23 100.0 735 12 BI762063 603049235
116 23 100.0 743 14 CD104670 AGENCOURT
117 23 100.0 744 12 BI597401 603243611
118 23 100.0 745 14 CB984378 AGENCOURT
119 23 100.0 751 12 BG759996 602733419
120 23 100.0 755 10 BF976108 602244951
121 23 100.0 756 12 BI835113 603087778
122 23 100.0 758 14 CB985816 AGENCOURT
123 23 100.0 761 12 BG541682 602571259
124 23 100.0 766 12 BG541205 602569986
125 23 100.0 766 12 BG570600 602591286
126 23 100.0 771 12 BI753761 603028285
127 23 100.0 776 12 BI227278 602948666
128 23 100.0 785 9 AV762204 AV762204
129 23 100.0 786 12 BG427527 602494551
130 23 100.0 789 12 BG538956 602568315
131 23 100.0 790 12 BI837330 603090674
132 23 100.0 791 13 BU596174 AGENCOURT
133 23 100.0 796 12 BG399870 602441984
134 23 100.0 796 14 CB997069 AGENCOURT
135 23 100.0 797 12 BG758487 602712686
136 23 100.0 799 9 AV752267 AV752267
137 23 100.0 802 10 BF238242 601811992
138 23 100.0 802 12 BG757550 602714723
139 23 100.0 803 12 BG759447 602711990
140 23 100.0 806 14 CB993192 AGENCOURT
141 23 100.0 809 14 CB992720 AGENCOURT
142 23 100.0 810 9 AV764215 AV764215
143 23 100.0 811 12 BI766898 603053138
144 23 100.0 817 12 BG758419 602712617
145 23 100.0 821 12 BI668794 603294742
146 23 100.0 821 12 BI755798 603024843
147 23 100.0 821 12 BI762843 603048376
148 23 100.0 822 12 BG341032 602462585
149 23 100.0 823 14 CB991431 AGENCOURT
150 23 100.0 826 12 BG431923 602498354

```

```

151 23 100.0 830 14 CB994002 AGENCOURT
152 23 100.0 831 10 BF525771 AGENCOURT
153 23 100.0 832 14 CB958376 AGENCOURT
154 23 100.0 837 12 BI517599 603042205
155 23 100.0 843 10 BF975896 602246288
156 23 100.0 843 12 BI261592 602953662
157 23 100.0 848 12 BG340681 602462262
158 23 100.0 851 12 BI769021 603058180
159 23 100.0 851 14 CD558780 AGENCOURT
160 23 100.0 854 12 BG756171 602713472
161 23 100.0 857 13 BQ890384 AGENCOURT
162 23 100.0 858 12 BI833342 603087948
163 23 100.0 859 12 BG546165 602573582
164 23 100.0 860 12 BG755507 602713882
165 23 100.0 864 10 BF663966 602145632
166 23 100.0 864 12 BI767351 603057683
167 23 100.0 872 12 BG542978 602569375
168 23 100.0 872 13 BQ721555 AGENCOURT
169 23 100.0 873 12 BG540219 602569108
170 23 100.0 874 12 BQ049475 AGENCOURT
171 23 100.0 875 12 BI821191 603035095
172 23 100.0 878 14 CB986657 AGENCOURT
173 23 100.0 879 13 BQ529093 AGENCOURT
174 23 100.0 880 9 AU139061 AU139061
175 23 100.0 881 14 CD558622 AGENCOURT
176 23 100.0 882 10 BZ662838 602144704
177 23 100.0 883 12 BG539370 602567984
178 23 100.0 884 12 BI545349 603187481
179 23 100.0 889 14 CD516536 AGENCOURT
180 23 100.0 890 12 BG754449 602710081

```

## ALIGNMENTS

```

RESULT 1
T91602 159 bp mrna linear EST 22-MAR-1995
LOCUS Y921b03.r1 Stragene lung (#937210) Homo sapiens cDNA clone
DEFINITION IMAGE:118349.5', similar to gb:K01506 HLA CLASS II
HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN); mRNA
sequence.
T91602
T91602.1 GI:723515
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (Bases 1 to 159)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissee, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, X., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Roifling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J.,
Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Maria, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 5 (9), 807-828 (1996)
JOURNAL 97044478
MEDLINE 8889549
PUBMED 8889549
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 114
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M3RP1
High quality sequence stop: 114.

```

```

FEATURES
  source
    Location/Qualifiers
      1..159
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="GDB:436638"
        /db_xref="taxon:9606"
        /clone="IMAGE:118349"
        /sex="male"
        /dev_stage="72 years"
        /lab_host="SOLR cells (kanamycin resistant)"
        /clone_lib="Stratagene lung (#937210)"
        /note="Organ: lung; Vector: pBluescript SK-; Site 1:
        EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
        Oligo dT, normal lung. Average insert size: 1.0 kb;
        Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG
        3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
      ORIGIN
        Query Match 100.0%; Score 23; DB 14; Length 159;
        Best Local Similarity 100.0%; Pred. No. 1.7;
        Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      QY 1 CGGACCATGTGTCAACTTATGCC 23
          |||||
      Db 52 CGGACCATGTGTCAACTTATGCC 74

      RESULT 2
      CD686854 168 bp mRNA linear EST 25-JUN-2003
      LOCUS
      DEFINITION EST33375 human nasopharynx Homo sapiens cDNA, mRNA sequence.
      ACCESSION CD686854
      VERSION CD686854.1 GI:32204158
      KEYWORDS EST.
      SOURCE
        Homo sapiens (human)
      ORGANISM
        Homo sapiens
      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      REFERENCE
        1. (bases 1 to 168)
        Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
        Zeng, Y.-X.
        Transcriptional Gene Expression Profile of Human Nasopharynx
        Unpublished (2003)
        Contact: Yixin Zeng
        Cancer Center
        Sun Yat-sen University
        651 Dongfeng Road East, Guangzhou 510060, China
        Tel: 86-1380-9770-743
        Fax: 86-20-8775-4506
        Email: yxzeng@gzsums.edu.cn.
      FEATURES
        source
          Location/Qualifiers
            1..168
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /tissue_type="normal nasopharynx"
              /clone_lib="human nasopharynx"
              /note="ESTs generated from a normal nasopharynx cDNA
              library from southern Chinese"
            ORIGIN
              Query Match 100.0%; Score 23; DB 14; Length 168;
              Best Local Similarity 100.0%; Pred. No. 1.7;
              Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

            QY 1 CGGACCATGTGTCAACTTATGCC 23
                |||||
            Db 123 CGGACCATGTGTCAACTTATGCC 145

            RESULT 3
            BG025808 177 bp mRNA linear EST 24-JAN-2001
            LOCUS

```

```

DEFINITION
  602274621F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4362701 5',
  mRNA sequence.
ACCESSION
  BG025808
  Version
  BG025808.1 GI:12412776
  KEYWORDS
  EST.
  SOURCE
    Homo sapiens (human)
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1. (bases 1 to 177)
    NIH-MGC http://img.nci.nih.gov/.
  AUTHORS
    National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE
    Unpublished (1999)
  JOURNAL
  COMMENT
    Contact: Robert Strausberg, Ph.D.
    Email: cgabs-r@mail.nih.gov
    Tissue procurement: Louis Staudt, M.D., Ph.D.
    cDNA Library Preparation: Life Technologies, Inc.
    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: LLAM10007 row: i column: 06
    High quality sequence stop: 138.
    Location/Qualifiers
      1..177
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4362701"
        /tissue_type="lymphoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_85"
        /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
        Average insert size 1.867 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH_MGC Library."
      ORIGIN
        Query Match 100.0%; Score 23; DB 10; Length 177;
        Best Local Similarity 100.0%; Pred. No. 1.7;
        Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      QY 1 CGGACCATGTGTCAACTTATGCC 23
          |||||
      Db 77 CGGACCATGTGTCAACTTATGCC 99

      RESULT 4
      BE140710/c 229 bp mRNA linear EST 21-JUN-2000
      LOCUS
      DEFINITION CM0-HT0016-140699-008 HT0016 Homo sapiens cDNA, mRNA sequence.
      ACCESSION BE140710
      VERSION BE140710.1 GI:8603431
      KEYWORDS EST.
      SOURCE
        Homo sapiens (human)
      ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      REFERENCE
        1. (bases 1 to 229)
        Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
        Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
        Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
        Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
        O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
        Simpson, A.J.
        Shotgun sequencing of the human transcriptome with ORF expressed
        sequence tags
        Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
      JOURNAL
      MEDLINE
      PUBMED
      COMMENT
        Contact: Simpson A.J.G.

```

Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0016-140>  
 699-008&t3=1999-06-14&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 7  
 High quality sequence stop: 229.  
 High quality sequence stop: 229.

## FEATURES

source

1. .229  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HTC016"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 23; DB 10; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACTTATGCC 23

Db 154 CGGACCATGTGTCACTTATGCC 132

## RESULT 5

BE159476  
 LOCUS BE159476 256 bp mRNA linear EST 21-JUN-2000  
 DEFINITION MR0-HT0407-100300-012-d07 HT0407 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BE159476  
 VERSION BE159476.1 GI:8622210  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 256)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.P., Matukuma, A., Bata, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MR0-HT0407-100>

300-012-d07&t3=2000-03-10&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 256.  
 Location/Qualifiers  
 1. .256

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0407"  
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 23; DB 10; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCACTTATGCC 23

Db 125 CGGACCATGTGTCACTTATGCC 147

## RESULT 6

BE242285

LOCUS

DEFINITION

BE242285 265 bp mRNA linear EST 03-OCT-2001  
 TCAAP1562 Pediatric acute myelogenous leukemia cell (FAB M1)  
 Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1562, mRNA

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@ccc.org

Citation: Carninci, P. and Hayashizaki, Y. High efficiency

full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M13 primer.

Location/Qualifiers

1. .265

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TCAAP1562"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="myeloid cell"

/dev\_stage="pediatric 6 years"

/lab\_host="DH10B"

/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project=TCAA"

/notes="Vector: lambda PSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'-GGAGACTCGAGCCGCCGAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer  
 [5'AGAGTCGATCCGGCGCGCAATTAATAT(C)3'].  
 Double-stranded cDNA was then digested with BamHI and XhoI  
 and directionally cloned into the BamHI and SalI sites of  
 lambda PSB vector. Library went through one round of  
 normalization. Library was constructed by Wei Yu at RIKEN  
 of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,  
 Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,  
 Schneider C, Hayashizaki Y, High efficiency selection of  
 full-length cDNA by improved biotinylated cap trapper.,  
 DNA Res 4: 1, 61-6, Feb 28, 1997)"

## ORIGIN

Query Match 100.0%; Score 23; DB 10; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 156 CGGACCATGTGTCAACTTATGCC 178

## RESULT 7

AA360953

LOCUS

DEFINITION

EST70157 T-cell lymphoma Homo sapiens cDNA 5' end similar to major  
 histocompatibility complex, class II antigen, alpha chain

(GB:X03100), mRNA sequence.

ACCESSION

AA360953

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,  
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,  
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,E., Fine,L.D.,  
 Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghagen,N.S.,  
 Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
 Morenc-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,  
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Lily,F.,  
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
 Dinke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,  
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,  
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Fields,C.,  
 Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseilte,W.A., Fields,C.,  
 Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns  
 based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

PUBMED

COMMENT

Other ESTs: THCI72266

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M3 Reverse.

Location/Qualifiers

1..289

/organism="Homo sapiens"

## FEATURES

source

/mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):165076"  
 /db\_xref="taxon:9606"  
 /cell\_type="T-lymphocyte"  
 /clone\_lib="T-cell lymphoma"  
 /notes="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI"

## ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 131 CGGACCATGTGTCAACTTATGCC 153

## RESULT 8

CD706205

LOCUS

DEFINITION

EST122732 human nasopharynx Homo sapiens cDNA, mRNA sequence.

ACCESSION

CD706205

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: YiXin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

Email: yxzeng@gzsums.edu.cn.

Location/Qualifiers

1..290

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="normal nasopharynx"

/clone\_lib="human nasopharynx"

/notes="ESTs generated from a normal nasopharynx cDNA

library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 169 CGGACCATGTGTCAACTTATGCC 191

## RESULT 9

CD694567

LOCUS

DEFINITION

EST11090 human nasopharynx Homo sapiens cDNA, mRNA sequence.

ACCESSION

CD694567

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 295)

AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.

TITLE Transcriptional Gene Expression Profile of Human Nasopharynx

JOURNAL Unpublished (2003)

COMMENT Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@gzsums.edu.cn.

FEATURES source  
1. .295  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 295;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||

Db 160 CGGACCATGTGTCAACTTATGCC 182  
|||||

RESULT 10  
AW404262 318 bp mRNA linear EST 16-FEB-2000  
LOCUS UI-HF-BL0-abg-b-07-0-UI.x1 NIH\_MGC\_37 Homo sapiens cDNA clone  
DEFINITION IMAGE.3057637 5', mRNA sequence.

ACCESSION AW404262  
VERSION AW404262.1 GI:6923319  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 318)  
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M.B. Soares Lab  
cDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:  
www-bio.lnl.gov/bbrp/image/image.html  
seq primer: M13 Forward.  
Location/Qualifiers  
1. .318  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3057637"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH MGC 37"  
/note="Vector: pTZ19-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(1.5-2.5kb). Directionally cloned. Cells provided by Louis  
M. Staudt, Ph.D. Library preparation by Maria de Fatima

## ORIGIN

Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

Query Match 100.0%; Score 23; DB 10; Length 318;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||

Db 62 CGGACCATGTGTCAACTTATGCC 84  
|||||

## RESULT 11

CD704658 334 bp mRNA linear EST 25-JUN-2003  
LOCUS EST21185 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
DEFINITION CD704658  
ACCESSION CD704658  
VERSION CD704658.1 GI:32235288  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 334)  
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.

Transcriptional Gene Expression Profile of Human Nasopharynx  
Unpublished (2003)  
COMMENT Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@gzsums.edu.cn.

## FEATURES source

1. .334  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 334;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23  
|||||

Db 219 CGGACCATGTGTCAACTTATGCC 241  
|||||

## RESULT 12

BM772902 335 bp mRNA linear EST 04-MAR-2002  
LOCUS K-BST0057161 S1SNUS Homo sapiens cDNA clone S1SNUS-8-C07 5', mRNA sequence.  
DEFINITION BM772902  
ACCESSION BM772902  
VERSION BM772902.1 GI:19102517  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 335)  
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.

21C Frontier Korean EST Project 2001  
Unpublished (2002)

## COMMENT

Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 8 row: C column: 07  
High quality sequence stop: 335.

## FEATURES

source

1. .335  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SISNU5-8-C07"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-5"  
/lab\_host="Top10F"  
/clone\_lib="SISNU5"  
/notes="Organ: Stomach; Vector: pONS; Site1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoRI  
- site by treatment of 14 RNA ligase and the first strand  
cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transfection of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

Query Match 100.0%; Score 23; DB 12; Length 335;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGGACCATGTGTCAACTTATGCC 23

Db 180 CGGACCATGTGTCAACTTATGCC 202

## RESULT 13

AA323639

LOCUS

DEFINITION  
EST26729 Cerebellum II Homo sapiens cDNA 5' end similar to major  
histocompatibility complex, class II antigen, alpha chain  
(GB:X03100), mRNA sequence.

AA323639

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 338)

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,

Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,

White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,

Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,

Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,

Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,B., Hinkle,P.S.Jr.,

Kelley,J.M., Kelley,J.C., Liu,J.-I., Marmaros,S.M., Merrick,J.M.,

Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,

Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,

Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,

Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,

## TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,  
Ke,W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,  
Raymond,L., Wei,X.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,  
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,  
Praser,C.M. and Venter,J.C.  
Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl), 3-174 (1995)  
96026280  
7566098  
Other ESTs: THCL172266  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tldb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

## FEATURES

source

1. .338  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):124435"  
/db\_xref="taxon:9606"  
/tissue\_type="cerebellum"  
/dev\_stage="adult"  
/clone\_lib="Cerebellum II"  
/notes="Organ: brain; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 100.0%; Score 23; DB 9; Length 338;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGGACCATGTGTCAACTTATGCC 23

Db 177 CGGACCATGTGTCAACTTATGCC 199

## RESULT 14

T94759

LOCUS

DEFINITION

T94759 366 bp mRNA linear EST 24-MAR-1995  
YE37f12.r1 Stratagene lung (#937210) Homo sapiens cDNA clone  
IMAGE:119951.5' similar to gb:K01506 HLA CLASS II  
HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN);, mRNA  
sequence.

T94759

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 366)

Hallier,L., Lennon,G., Becker,M., Bonaudo,M.F., Chiapelli,B.,

Chissoe,S., Dietrich,N., DuBuque,T., Pavello,A., Gish,W.,

Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,

Mardis,E., Moore,B., Mortis,M., Parsons,J., Prange,C., Rifkin,L.,

Rolling,T., Scellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,

Trevasakis,E., Underwood,K., Woldmann,P., Waterston,R., Wilson,R.,

and Marra,M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (5), 807-828 (1996)

97044478

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Wilson RK

Washington University School of Medicine



4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286-1800

Fax: 314 286-1810

Email: est@watson.wustl.edu

Insert Size: 1152

High quality sequence stops: 251 Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 1152 Std Error: 0.00

Seq primer: M13RPI

High quality sequence stop: 251.

#### FEATURES

source  
1..366  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:488240"  
/db\_xref="taxon:9606"  
/clone="IMAGE:119951"  
/sex="male"  
/dev\_stage="72 years"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/clone\_lib="Stratagene lung (#937210)"  
/note="Organ: lung; Vector: pBluescript SK-; Site: 1:  
EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. normal lung. Average insert size: 1.0 kb;  
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGACGAG  
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

#### ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 366;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 50 CGGACCATGTGTCAACTTATGCC 72

#### RESULT 15

LOCUS

CB267061 1005967 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens

CDNA 5'; mRNA sequence.

CB267061

CB267061.1 GI:28441647

EST.

source

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 387)

Yang, R.-Z., Shuldiner, A. and Gong, D.-W.

EST analysis of human adipose gene expression

Unpublished (2002)

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, HH497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGAGCGCGCATCTGTGTGT

BACKWARD: ATACGACTCATTATAGGGGATTTG

Seq primer: GTTGTACCGGATTC.

location/Qualifiers

1..387

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/sex="Male and Female"

/tissue type="Adipose"

/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

#### ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 387;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 131 CGGACCATGTGTCAACTTATGCC 153

#### RESULT 16

LOCUS

CD709023

DEFINITION

EST25550 human nasopharynx Homo sapiens cDNA, mRNA sequence.

CD709023

VERSION

CD709023.1 GI:32239653

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 390)

Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and

Zeng, Y.-X.

Transcriptional Gene Expression Profile of Human Nasopharynx

Unpublished (2003)

Contact: Yixin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

Email: yxzeng@gzsums.edu.cn.

location/Qualifiers

1..390

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue type="normal nasopharynx"

/clone\_lib="human nasopharynx"

/note="ESTs generated from a normal nasopharynx cDNA

library from southern Chinese"

#### ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 390;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCATGTGTCAACTTATGCC 23

Db 189 CGGACCATGTGTCAACTTATGCC 211

#### RESULT 17

LOCUS

AA244273

DEFINITION

nc06d06.t1 NCI CGAP P11 Homo sapiens cDNA clone IMAGE:1007339

similar to gb:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)

ALPHA CHAIN (HUMAN);, mRNA sequence.

AA244273

VERSION

AA244273.1 GI:1875023

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 405)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

#### FEATURES

source  
1..387  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="Male and Female"  
/tissue type="Adipose"  
/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

```

COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: csapbs-remail.nih.gov
              Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
              M.D., Michael Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: David B. Krizman, Ph.D.
              cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E.S. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Insert Length: 673      Std Error: 0.00
              Seq primer: -28ml3 rev1 ET from Amersham
              High quality sequence stop: 371.
              Location/Qualifiers
FEATURES
    source
    1..408
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="S20T665307-44-C11"
    /sex="M"
    /lab_host="Top10F"
    /clone_lib="S20T665307"
    /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
    Site 2: NotI; The poly (A) + RNA was dephosphorylated with
    bacterial alkaline phosphatase (BAP) and then decapped
    with tobacco acid pyrophosphatase (TAP). The decapped
    intact mRNA was ligated with DNA-RNA linker including EcoR
    I site by treatment of T4 RNA ligase and the first strand
    cDNA was synthesized from oligo dt-selected mRNA by
    priming with dt-tailed vector. The dt-tailed vector was
    adjusted to have about 60nt. The cDNA vector was
    circularized with E. coli DNA ligase after digestion of
    EcoRI which site is also included in vector. An RNA strand
    converted to a DNA strand by Okayama-Berg method. The
    obtained cDNA vectors were used for transformation of
    competent cells E. coli Top10F by electroporation method.
    The cDNA libraries constructed by this method are
    full-length enriched cDNA library."
ORIGIN
    Query Match      100.0%; Score 23; DB 12; Length 408;
    Best Local Similarity 100.0%; Pred. No. 2.1;
    Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 CGGACCATGTGTCAACTTATGCC 23
    Db 151 CGGACCATGTGTCAACTTATGCC 173

RESULT 19
AA838010
LOCUS      AA838010      415 bp      mRNA      linear      EST 26-FEB-1998
DEFINITION oes9d07.s1 NCI CGAP Col2 Homo sapiens cDNA clone IMAGE:1418797
            similar to gb:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)
            ALPHA CHAIN (HUMAN); contains Alu repetitive element;; mRNA
            sequence.
ACCESSION  AA838010
VERSION     AA838010.1      GI:2913667
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 415)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: csapbs-remail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Stratagene, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E.S. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 301.
            Location/Qualifiers
FEATURES
    source
    1..415
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAG:1418797"
    /sex="mixed"

```

```

/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Col2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
GAATTCGCGACGAG 3' 3' adaptor sequence: 5'
CTCGAGTTCTTTTCTTTT 3' Average insert size: 1.2 kb."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 23; DB 9; Length 415;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 90 CGGACCATGTGTCAACTTATGCC 112

RESULT 20
LOCUS BF378520 418 bp mRNA linear EST 27-NOV-2000
DEFINITION QV1-UM0036-200300-115-a07 UM0036 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF378520
VERSION BF378520.1 GI:11367658
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 418)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV1&t2=QV1-UM0036-
200300-115-a07&t3=2000-03-20&t4=1)
Seq primer: puc 18 forward.
Location/Qualifiers
1..418
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0036"
/note="Organ: uterus; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORF-ES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 23; DB 10; Length 418;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 33 CGGACCATGTGTCAACTTATGCC 55

RESULT 21
LOCUS C03540 422 bp mRNA linear EST 30-JUL-1996
DEFINITION C03540 Human heart cDNA (YNakamura) Homo sapiens cDNA clone
3NHC1710, mRNA sequence.
ACCESSION C03540
VERSION C03540.1 GI:1466791
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 422)
Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
Nakamura,Y.
Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
96299762
8661126
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..422
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="3NHC1710"
/dev_stage="adult"
/clone_lib="Human heart cDNA (YNakamura)"
/note="Organ: heart; normalized directionally cloned cDNA
from adult heart"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 23; DB 13; Length 422;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCATGTGTCAACTTATGCC 23
Db 175 CGGACCATGTGTCAACTTATGCC 197

RESULT 22
LOCUS CD102141 423 bp mRNA linear EST 15-MAY-2003
DEFINITION AGENCOURT 14007825 NIH MGC 186 Homo sapiens cDNA clone
IMAGE:30370583 5', mRNA sequence.
ACCESSION CD102141
VERSION CD102141.1 GI:30755315
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 604.235 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819B-35

Perfect score: 20

Sequence: 1 aggaagagagagcgctgc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB Length	ID	Description
1	18.4	92.0	139056	2	AC141691	AC141691 Apis mell
2	18.4	92.0	188090	9	AC011625	AC011625 Homo sapi
3	18.4	92.0	219368	10	AC098721	AC098721 Mus muscu
4	18	90.0	133984	9	AL512306	AL512306 Human DNA
5	18	90.0	153023	2	AC021462	AC021462 Homo sapi
6	17.4	87.0	28134	3	CER186	278016 Caenorhabdi
7	17	85.0	124479	2	AC020966	AC020966 Mus muscu
8	16.8	84.0	1014	1	TTY14655	Y14655 Thermoprote
9	16.8	84.0	3312	5	CHKPLB	MS9039 Chicken car
10	16.8	84.0	3423	5	CHKPLB2	MS9038 Chicken pho
11	16.8	84.0	3491	8	AK065803	AK065803 Oryza sat
12	16.8	84.0	6176	9	P247811S12	AF238295 Homo sapi
13	16.8	84.0	28396	3	U41264	U41264 Caenorhabdi
14	16.8	84.0	33010	5	AL592204	AL592204 Zebrafish
15	16.8	84.0	52423	2	AC020395	AC020395 Drosophila
16	16.8	84.0	71503	9	AF219991	AF219991 Homo sapi
17	16.8	84.0	80423	3	DROABDB	L07835 Drosophila
18	16.8	84.0	118948	2	AP003226	AP003226 Oryza sat
19	16.8	84.0	121088	8	AP004326	AP004326 Oryza sat
20	16.8	84.0	125364	4	AC139728	AC139728 Carolina
21	16.8	84.0	132060	2	AC138555	AC138555 Carolinia
22	16.8	84.0	132900	8	AC113948	AC113948 Oryza sat
23	16.8	84.0	136548	8	AP004368	AP004368 Oryza sat
24	16.8	84.0	142472	9	AC142304	AC142304 Pan trogl
25	16.8	84.0	144596	8	AP004574	AP004574 Oryza sat
26	16.8	84.0	148012	2	AC138995	AC138995 Homo sapi
27	16.8	84.0	149061	9	AC098805	AC098805 Homo sapi
28	16.8	84.0	151071	8	AP003449	AP003449 Oryza sat
29	16.8	84.0	151249	2	AC135250	AC135250 Rattus no
30	16.8	84.0	151334	2	BX571897	BX571897 Danio rer
31	16.8	84.0	152316	10	AC007844	AC007844 Mus muscu
32	16.8	84.0	152802	8	AP005491	AP005491 Oryza sat
33	16.8	84.0	156173	9	AC109792	AC109792 Bos tauru
34	16.8	84.0	157358	9	AC009163	AC009163 Homo sapi
35	16.8	84.0	160613	2	AC134969	AC134969 Homo sapi
36	16.8	84.0	165165	3	AC095018	AC095018 Drosophila
37	16.8	84.0	175335	3	AC091636	AC091636 Drosophila
38	16.8	84.0	176552	9	AC106820	AC106820 Homo sapi
39	16.8	84.0	179886	10	AC114826	AC114826 Mus muscu
40	16.8	84.0	183037	2	AC136111	AC136111 Rattus no
41	16.8	84.0	183538	2	AC147272	AC147272 Pan trogl
42	16.8	84.0	194361	8	AC092389	AC092389 Oryza sat
43	16.8	84.0	194832	9	AC092587	AC092587 Homo sapi
44	16.8	84.0	201853	9	AL591424	AL591424 Human DNA
45	16.8	84.0	203284	2	AC110130	AC110130 Rattus no
46	16.8	84.0	205268	2	AC146518	AC146518 Homo sapi
47	16.8	84.0	212221	2	AC136126	AC136126 Rattus no
48	16.8	84.0	227877	2	AC109725	AC109725 Rattus no
49	16.8	84.0	238544	3	AE003715	AE003715 Drosophila
50	16.8	84.0	239161	2	AC128213	AC128213 Rattus no
51	16.8	84.0	232945	2	AC103156	AC103156 Rattus no
52	16.8	84.0	238596	2	AC119587	AC119587 Rattus no
53	16.8	84.0	239080	2	AC111635	AC111635 Rattus no
54	16.8	84.0	240550	2	AC111385	AC111385 Rattus no
55	16.8	84.0	241178	2	AC130508	AC130508 Rattus no
56	16.8	84.0	244843	2	AC107162	AC107162 Rattus no
57	16.8	84.0	252978	2	AC126983	AC126983 Rattus no
58	16.8	84.0	266661	2	AC126523	AC126523 Rattus no
59	16.8	84.0	295712	2	AC120456	AC120456 Rattus no
60	16.8	84.0	296282	2	AC111857	AC111857 Rattus no
61	16.8	84.0	300029	8	AE017114	AE017114 Oryza sat
62	16.8	84.0	316828	2	AC135386	AC135386 Rattus no
63	16.8	84.0	338234	3	DMU1961	U31961 Drosophila
64	16.4	82.0	37068	8	AC067938	AC067938 Neurospor
65	16.4	82.0	43556	8	AC067937	AC067937 Neurospor

66 16.4 82.0 59429 2 AC091176 Homo sapi  
67 16.4 82.0 63853 2 AC087663 Homo sapi  
68 16.4 82.0 68217 9 AL590664 Human DNA  
69 16.4 82.0 69208 2 AC020466 Drosophila  
70 16.4 82.0 74512 8 AC100521 Mus muscu  
71 16.4 82.0 93593 8 AF263283 Filobasid  
72 16.4 82.0 93979 8 AC068564 Filobasid  
73 16.4 82.0 95663 9 AC010247 Homo sapi  
74 16.4 82.0 128323 8 AC132215 Genomic s  
75 16.4 82.0 159330 2 AC138757 Homo sapi  
76 16.4 82.0 167108 2 AC068283 Homo sapi  
77 16.4 82.0 175118 3 AC010842 Drosophila  
78 16.4 82.0 205488 2 AC115975 Mus muscu  
79 16.4 82.0 220371 2 AC109966 Rattus no  
80 16.4 82.0 240051 2 AC122326 Mus muscu  
81 16.4 82.0 268335 2 AC126066 Rattus no  
82 16.4 82.0 295225 3 AC003461 Drosophila  
83 16.4 82.0 303862 1 AC017215 Geobacter  
84 16 80.0 155164 9 AC093182 Homo sapi  
85 16 80.0 176209 9 AC009518 Homo sapi  
86 16 80.0 195811 2 AC147038 Pan trogl  
87 16 80.0 213604 2 AC100345 Mus muscu  
88 16 80.0 236527 2 AC147103 Pan trogl  
89 16 80.0 274669 2 BX005328 Drosophila  
90 15.8 79.0 668 8 CNS01A2X  
91 15.8 79.0 720 8 CNS01944  
92 15.8 79.0 780 8 CNS0194P  
93 15.8 79.0 780 8 CNS01B3CD  
94 15.8 79.0 869 3 SP22B1G  
95 15.8 79.0 951 8 AKI03638  
96 15.8 79.0 1092 9 HSA403950  
97 15.8 79.0 1362 6 AX431819 Sequence  
98 15.8 79.0 1770 10 BC013462 Mus muscu  
99 15.8 79.0 2000 6 AX566405 Sequence  
100 15.8 79.0 2000 6 AX566673 Sequence  
101 15.8 79.0 3059 8 AK069509 Oryza sat  
102 15.8 79.0 3213 9 HSEXM50R  
103 15.8 79.0 3770 8 AKI21357 Oryza sat  
104 15.8 79.0 3882 10 BC057380 Mus muscu  
105 15.8 79.0 6574 3 WSPRTANSE  
106 15.8 79.0 7690 6 AX346024 Sequence  
107 15.8 79.0 7950 1 AF016298 Rhodobact  
108 15.8 79.0 11002 1 AF117827 Methyloco  
109 15.8 79.0 14924 6 AX281280 Sequence  
110 15.8 79.0 14924 6 AX345127 Sequence  
111 15.8 79.0 39351 3 AC105441 Leishmani  
112 15.8 79.0 39349 3 AC009782 Leishmani  
113 15.8 79.0 40679 9 AC002052 Homo sapi  
114 15.8 79.0 58020 2 AP002772 Homo sapi  
115 15.8 79.0 58852 2 AC036190 Homo sapi  
116 15.8 79.0 61076 2 AC104860 Mus muscu  
117 15.8 79.0 64401 2 AC131908 Homo sapi  
118 15.8 79.0 64523 2 AC100987 Homo sapi  
119 15.8 79.0 67427 2 AC101742 Mus muscu  
120 15.8 79.0 70752 2 AC113148 Homo sapi  
121 15.8 79.0 71241 2 AC102061 Mus muscu  
122 15.8 79.0 71241 2 AC102061 Mus muscu  
123 15.8 79.0 72400 8 AP005524 Oryza sat  
124 15.8 79.0 78770 2 AC034230 Homo sapi  
125 15.8 79.0 80465 2 AL359386 Homo sapi  
126 15.8 79.0 82270 9 AL354879 Human DNA  
127 15.8 79.0 86722 8 AB012245 Arabidops  
128 15.8 79.0 91826 9 AB045320 Homo sapi  
129 15.8 79.0 92455 2 AC034136 Homo sapi  
130 15.8 79.0 92510 9 HS390B3  
131 15.8 79.0 93240 2 AC007586 Drosophila  
132 15.8 79.0 95597 2 AC014032 Drosophila  
133 15.8 79.0 106256 3 AC108135 Leishmani  
134 15.8 79.0 108539 2 AC134299 Homo sapi  
135 15.8 79.0 108553 8 AF527809 Sorghum b  
136 15.8 79.0 110000 2 AC112799 Rattus no  
137 15.8 79.0 110000 2 AC12799-1  
138 15.8 79.0 110000 2 IMFLCHR18\_08

139 15.8 79.0 110000 2 IMFLCHR36\_28  
140 15.8 79.0 111111 10 AC130816 Mus muscu  
141 15.8 79.0 112630 2 AC025461 Homo sapi  
142 15.8 79.0 119171 2 BX247904 Dancio rer  
143 15.8 79.0 121041 9 AC008042 Homo sapi  
144 15.8 79.0 122940 8 AP004054 Oryza sat  
145 15.8 79.0 123013 2 AC010005 Drosophila  
146 15.8 79.0 127218 2 AC141738 Apis mell  
147 15.8 79.0 127675 9 AC146221 Pan trogl  
148 15.8 79.0 130027 9 AC009773 Homo sapi  
149 15.8 79.0 130193 8 AC138004 Oryza sat  
150 15.8 79.0 132910 3 AC014319 Drosophila  
151 15.8 79.0 133713 3 AC087838 Leishmani  
152 15.8 79.0 135940 9 AC119406 Trypanoso  
153 15.8 79.0 140100 9 AC016778 Homo sapi  
154 15.8 79.0 141489 2 AC141017 Rattus no  
155 15.8 79.0 142195 2 AC010669 Drosophila  
156 15.8 79.0 146558 2 AP004868 Oryza sat  
157 15.8 79.0 146717 3 AC097726 Drosophila  
158 15.8 79.0 149041 9 AC044782 Homo sapi  
159 15.8 79.0 149716 2 AC119380 Lemur cat  
160 15.8 79.0 150377 2 AC093405 Oryza sat  
161 15.8 79.0 150995 8 AC104615 Oryza sat  
162 15.8 79.0 151967 10 AL805912 Mouse DNA  
163 15.8 79.0 153149 9 AL354819 Human DNA  
164 15.8 79.0 153508 2 AC141790 Apis mell  
165 15.8 79.0 154954 9 AC022507 Homo sapi  
166 15.8 79.0 156221 2 AC134963 Canis fam  
167 15.8 79.0 159156 2 AC048350 Homo sapi  
168 15.8 79.0 160252 2 AL390250 Homo sapi  
169 15.8 79.0 160785 9 AC066583 Homo sapi  
170 15.8 79.0 161741 8 AC120535 Oryza sat  
171 15.8 79.0 161852 2 AC135690 Rattus no  
172 15.8 79.0 162083 9 AL591069 Human DNA  
173 15.8 79.0 163043 2 AC124654 Homo sapi  
174 15.8 79.0 163066 2 AC116506 Mus muscu  
175 15.8 79.0 163475 10 AC026762 Mus muscu  
176 15.8 79.0 164812 9 AC078820 Homo sapi  
177 15.8 79.0 165594 2 AC023317 Homo sapi  
178 15.8 79.0 167408 2 AC025526 Homo sapi  
179 15.8 79.0 167711 2 AC079166 Homo sapi  
180 15.8 79.0 168990 2 AC122516 Mus muscu

## ALIGNMENTS

RESULT 1  
AC141691

LOCUS  
DEFINITION  
AC141691 linear HTG 19-MAR-2003  
Apis mellifera clone CH224-5515, WORKING DRAFT SEQUENCE, 30  
unordered pieces.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AC141691.1 GI:29123875  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Apis mellifera (honeybee)  
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
Apidae; Apis.

REFERENCE  
AUTHORS

1 (bases 1 to 139056)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Albrooks,S.L., Amarutunge,H.C., Are,J.R., Ayale,M., Banks,P.,  
Barbieri,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,J.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,X.,

Falls, T., Ferraguto, D., Flagg N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrrell, J. H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, B., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratochvic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louissege, H., Lozaro, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mathewine, B., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miaz, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Ogarunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pul, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, T., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taylor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmari, K., Vasquez, L., Vera, V., Vallalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodden, S., Wcley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 139056)  
Worley, K.C.  
Direct Submission  
Submitted (19-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: AMGP
Center clone name: GH224-5515
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator; Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 125289 bases at least Q40
Consensus quality: 128700 bases at least Q30
Consensus quality: 130487 bases at least Q20
Estimated insert size: 129698; sum-of-contigs estimation
Quality coverage 3x in Q20 bases; sum-of-contigs estimation

```

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.bgsb.cmc.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.bgsb.cmc.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a working draft, sequence is currently  
 \* consists of 30 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* \*

*	1	1585:	config	of	1555	bp	in	length
*		1556	1655:	gap	of	unknown	length	
*	1556	3079:	config	cf	1424	bp	in	length
*	3080	3179:	gap	of	unknown	length		
*	3180	4896:	config	of	1717	bp	in	length
*	4897	4996:	gap	of	unknown	length		
*	4997	6870:	config	of	1874	bp	in	length
*	6871	9284:	gap	of	unknown	length		
*	9284:	config	of	1314	bp	in	length	

*	8285	8384: gap of unknown length
*	8385	10898: contig of 2514 bp in length
*	10898	10898: gap of unknown length
*	10999	12432: contig of 1434 bp in length
*	12433	12533: gap of unknown length
*	12533	14591: contig of 2059 bp in length
*	12533	14591: contig of 2059 bp in length
*	14592	14691: gap of unknown length
*	14592	16011: contig of 1324 bp in length
*	16011	16113: gap of unknown length
*	16116	18836: contig of 2721 bp in length
*	18837	18936: gap of unknown length
*	18937	20628: contig of 1692 bp in length
*	20629	20728: gap of unknown length
*	20729	236113: contig of 2885 bp in length
*	236114	237113: gap of unknown length
*	237114	26592: contig of 2879 bp in length
*	26593	26692: gap of unknown length
*	26693	29806: contig of 3114 bp in length
*	29807	29906: gap of unknown length
*	29907	34149: contig of 4243 bp in length
*	34150	34249: gap of unknown length
*	34250	37904: contig of 3655 bp in length
*	37905	38004: gap of unknown length
*	38005	40918: contig of 2914 bp in length
*	40919	41018: gap of unknown length
*	41019	45267: contig of 4249 bp in length
*	45268	45367: gap of unknown length
*	45368	50580: contig of 5213 bp in length
*	50581	50680: gap of unknown length
*	50681	55750: contig of 5070 bp in length
*	55751	55850: gap of unknown length
*	55851	60417: contig of 4567 bp in length
*	60418	60517: gap of unknown length
*	60518	65870: contig of 5353 bp in length
*	65871	65970: gap of unknown length
*	65971	70676: contig of 4706 bp in length
*	70677	70776: gap of unknown length
*	70777	77788: contig of 7012 bp in length
*	77789	77888: gap of unknown length
*	77889	82008: contig of 4120 bp in length
*	82009	82108: gap of unknown length
*	82109	88870: contig of 6762 bp in length
*	88871	88970: gap of unknown length
*	88971	98432: contig of 9462 bp in length
*	98433	98532: gap of unknown length
*	98533	108199: contig of 9684 bp in length
*	108197	108296: gap of unknown length
*	108297	123260: contig of 14064 bp in length
*	123261	124460: gap of unknown length
*	124461	139566: contig of 16596 bp in length.

```

location/Qualifiers
  . 139056
  /organism="Apis mellifera"
  /mol_type="Genomic DNA"
  /db_xref="taxon:7460"
  /clone="CH224-5S15"

```

```

          92.0%; Score 18.4; DB 2; Length 1
1 Similarity 95.0%; Pred. NC. 73;
19; Conservative 0; Mismatches 1; Indels
1 AGGAAAGGAGAGCGTCGTC 20
      |||||
      |||||CGCCCTCCTGC 3730

```

AC011625	188390 bp	DNA	linear
Homo sapiens clone RP11-341C17 from 7q31,			complete
AC011625			
AC011625.2	GI:6539285		

**KEYWORDS**  
**SOURCE** HTG.  
**ORGANISM** Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
**AUTHORS** 1 (bases 1 to 188090)  
 Kaul,R.K., Xu,J., Wong,G.K.-S., Magnus,C.L., Green,E.D., Green,P.  
 and Olson,M.V.  
**TITLE** Large-scale MCD Mapping and Sequencing of Human Chromosome 7  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 188090)  
 Bubb,K.L., Desmarais,C.L., Ramsey,S.A. and Hubley,R.M.  
**AUTHORS** Direct Submission  
**TITLE** Submitted (08-Oct-1999) Human Genome Center, University of  
 Washington, Box 352145, Seattle, WA 98195, USA  
**JOURNAL** 3 (bases 1 to 188090)  
 Kaul,R.K. and Desmarais,C.L.  
**REFERENCE** Direct Submission  
**AUTHORS** Submitted (08-Dec-1999) Human Genome Center, University of  
 Washington, Box 352145, Seattle, WA 98195, USA  
**TITLE** On Dec 8, 1999 this sequence version replaced gi:6016750.  
**JOURNAL**  
**COMMENT**

----- Genome Center:  
 University of Washington Genome Center  
 Center code: UWGC  
 Web site: <http://genome.washington.edu>  
 Contact: [uwgchelp@u.washington.edu](mailto:uwgchelp@u.washington.edu)  
 ----- Project Information  
 Center project name: chr-7  
 Center clone name: djs301 (RP11-341C17)  
 ----- Summary Statistics  
 Sequencing vector: M13; 100% of reads  
 Chemistry: Dye-primer Bodipy; 93% of reads Chemistry:  
 Dye-terminator Big Dye; 7% of reads  
 Assembly program: Phrap; version 0.990319  
 Insert size: 188 094; sum-of-contigs  
 Quality coverage: 7.78X in Q20 bases; sum-of-contigs  
 -----  
 Overlapping Sequences:  
 5': UWGC:djs302  
 3': UWGC:djs380  
 -----  
 Sequence Quality Assessment:  
 This entry has been annotated with sequence quality  
 estimates computed by the Phrap assembly program.  
 All manually edited bases have been reduced to quality zero.  
 Quality levels above 40 are expected to have less than  
 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the  
 GenBank flat file format but are available as part  
 of this entry's ASN.1 file.  
 -----  
 Double stranded (DS) coverage: 70.4%  
 DS or two chemistry coverage: 100.0%  
 Single stranded regions: 0  
 -----  
 Sequence Validation:  
 This sequence has been validated by Multiple Complete Digest  
 fingerprinting. Comparison of the experimentally derived digest  
 fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and  
 vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-600 bp)  
 are not resolved in the fingerprint and hence do not appear  
 in the table. There are no significant remaining discrepancies  
 between the experimental and predicted values. Uniquely ordered  
 fragments are separated by dashed lines.  
 -----  

FP	Seq	FP	Seq	FP	Seq
2095.00	2061.00	8947.00	8687.00	729.00	732.00
12518.00	12347.00	844.00	812.00	1463.00	1480.00

1524.00 1500.00 6169.00 6048.00 7052.00 6870.00  
 3776.00 3636.00 2207.00 2144.00 3731.00 3659.00  
 5704.00 5638.00  
 6982.00 6968.00  
 -----  
 -----

## FEATURES

source Location/Qualifiers  
 1. 188090  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7q31"  
 /clones="RP11-341C17 or UWGC:djs301"  
 /cell\_line="Male Blood"  
 /clone\_lib="RPC-11 Human Male BAC Library"  
 444. 750  
 /rpt\_family="Alu"  
 891. 963  
 /standard\_name="sWS2952"  
 /note="Genbank Accession: G13043"  
 complement(2131. 2409)  
 /rpt\_family="Alu"  
 complement(8590. 8831)  
 /rpt\_family="Alu"  
 10548. 10634  
 /rpt\_family="Alu"  
 11525. 11795  
 /rpt\_family="Alu"  
 18103. 18403  
 /rpt\_family="Alu"  
 complement(27089. 27369)  
 repeat\_region 92.0%; Score 18.4; DB 9; Length 188090;  
 Best Local Similarity 95.0%; Pred.No. 73;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGGAAGAGAGAGCGGTCGTC 20  
 |||||  
 Db 80784 AGGAAGAGAGAGCGGTCGTC 80765

RESULT 3  
 AC098721/c LOCUS 219368 bp DNA linear ROD 31-OCT-2003  
 DEFINITION Mus musculus BAC clone RP23-2023 from 2, complete sequence.  
 AC098721 ACCESSION  
 AC098721.3 GI:19909472  
 HTG. VERSION  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 219368)  
 REFERENCE  
 AUTHORS Armstrong,J., Abbott,A. and Creason,K.

The sequence of Mus musculus BAC clone RP23-2023  
 Unpublished (2001)  
 2 (bases 1 to 219368)  
 WILSON,R.

Sequencing of Mus musculus  
 Unpublished (2001)  
 3 (bases 1 to 219368)  
 McPherson,J.D. and Waterston,R.H.

Direct Submission  
 Submitted (31-OCT-2001) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 219368)  
 McPherson,J.D. and Waterston,R.H.

Direct Submission  
 Submitted (03-APR-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 219368)  
 McPherson,J.D. and Waterston,R.H.

Direct Submission  
 Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 219368)  
 Wilson,R.

Direct Submission  
 Submitted (31-OCT-2003) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Apr 3, 2002 this sequence version replaced gi17017625.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

----- Summary Statistics

----- Center project name: M\_BA0002023

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,  
 Department of Genetics, Washington University, St. Louis MO. For  
 additional information about the map position of this sequence, see  
 http://genome.wustl.edu

## SOURCE INFORMATION:

The RPI-23 BAC Library has been constructed by Kazutoyo Oseegawa  
 and Minako Iateno in the laboratory of Pieter de Jong  
 (http://www.chori.org) from female C57BL/6J mouse kidney and/or  
 brain genomic DNA. The clone and detailed information can be  
 obtained from Research Genetics, Inc. (http://www.resgen.com) or  
 Pieter de Jong and coworkers at http://www.chori.org

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

1. 219368  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /map="2"  
 /clone="RP23-2023"

## FEATURES

source



```

repeat_region /clone_lib="RPC1-23"
repeat_region 4..99
repeat_region /rpt_family="Alu"
repeat_region 1931..2204
repeat_region /rpt_family="L1"
repeat_region 3672..3749
repeat_region /rpt_family="Alu"
repeat_region 3832..3971
repeat_region /rpt_family="Alu"
repeat_region 6380..6576
repeat_region /rpt_family="B2"
repeat_region 7101..7291
repeat_region /rpt_family="B2"
repeat_region 9747..9855
repeat_region /rpt_family="L1"
repeat_region 9914..10361
repeat_region /rpt_family="L1"
repeat_region 17647..17740
repeat_region /rpt_family="MIR"
repeat_region 19477..19723
repeat_region /rpt_family="B4"
repeat_region 19721..19824
repeat_region /rpt_family="B4"
repeat_region 20827..21403
repeat_region /rpt_family="L1"
repeat_region 21442..22029
repeat_region /rpt_family="L1"
repeat_region 23389..23433
misc_feature /note="Unresolved homopolymeric repeat."
repeat_region 23652..23763
repeat_region /rpt_family="L1"
repeat_region 23791..23925
repeat_region /rpt_family="Alu"
repeat_region 24038..24740
repeat_region /rpt_family="L1"
repeat_region 24731..25012
repeat_region /rpt_family="L1"
repeat_region 25372..25932
repeat_region /rpt_family="L1"
repeat_region 25926..26093
repeat_region /rpt_family="L1"
repeat_region 26350..26745
repeat_region /rpt_family="MaLR"
repeat_region 26896..27019
repeat_region /rpt_family="ERVX"
repeat_region 26993..27010
repeat_region /rpt_family="B4"
repeat_region 27011..27095
repeat_region /rpt_family="Alu"
repeat_region 27128..27495
repeat_region /rpt_family="ERVX"
repeat_region 27842..28110
repeat_region /rpt_family="MaLR"
repeat_region 28214..28446
repeat_region /rpt_family="L1"
repeat_region 30377..30522
repeat_region /rpt_family="Alu"
repeat_region 31240..31636
repeat_region /rpt_family="L1"
repeat_region 31631..31972
repeat_region /rpt_family="L1"
repeat_region 32577..32784
repeat_region /rpt_family="B2"
repeat_region 33405..33511
repeat_region /rpt_family="L1"
repeat_region 34749..35091
repeat_region /rpt_family="L1"
repeat_region 36138..36280
repeat_region /rpt_family="ERVX"
repeat_region 37095..37481
repeat_region /rpt_family="B2"
repeat_region 37875..37968
repeat_region /rpt_family="MIR"

repeat_region 44056..44148
repeat_region /rpt_family="Alu"
repeat_region 45143..45273
repeat_region /rpt_family="MER1_type"
repeat_region 45296..45504
repeat_region /rpt_family="MaLR"
repeat_region 45546..45954
repeat_region /rpt_family="L1"
repeat_region 45982..47257
repeat_region /rpt_family="MER1_type"
repeat_region 47306..47380
repeat_region /rpt_family="tRNA-Ala-GCY"
repeat_region 47394..47511
repeat_region /rpt_family="MER1_type"
repeat_region 49498..49991
repeat_region /rpt_family="ERVX"
repeat_region 50037..50184
repeat_region /rpt_family="Alu"
repeat_region 52307..53184
repeat_region /rpt_family="L1"
repeat_region 53651..53837
repeat_region /rpt_family="L1"
repeat_region 53848..53958
repeat_region /rpt_family="5S"
repeat_region 53985..54117
repeat_region /rpt_family="L1"
repeat_region 54119..54567
repeat_region /rpt_family="RMR6B"
repeat_region 55344..55466
repeat_region /rpt_family="ERVX"
repeat_region 55852..56355
repeat_region /rpt_family="L1"
repeat_region 56390..56440
repeat_region /rpt_family="Alu"
repeat_region 58409..59200
repeat_region /rpt_family="ERVX"
repeat_region 59202..60566
repeat_region /rpt_family="L1"
repeat_region 60573..60633
repeat_region /rpt_family="ERV1"
repeat_region 60753..61495
repeat_region /rpt_family="ERVX"
repeat_region 61705..61913

Query Match 92.0%; Score 18.4; DB 10; Length 219368;
Best Local Similarity 95.0%; Pred.No.73;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 3;

Qy 1 AGGAAGGAGGAGCGTCGTC 20
Db 89436 AGGAAGGAGGAGCGTCGTC 89417

RESULT 4
AL512306/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-430C7 on chromosome 1, complete
sequence.
ACCESSION AL512306
VERSION AL512306.16 GI:18491332
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 133984)
AUTHORS Bray-Allen,S.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:18477311.
During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submissiion corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP11-430C7 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.choxi.org/bacpac/home.htm> VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-430C7. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-430C7 is at 133984 in this sequence. The true left end of clone RP11-739N20 is at 97715 in this sequence. The true right end of clone RP11-2317 is at 2000 in this sequence.

```
FEATURES             Location/Qualifiers
     source            1..133984
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-430C7"
                     /clone_lib="RPCI-11.2"
     misc_feature      123904..123933
                     /note="Single clone region. Reads generated from a
                     transposon library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
```

## ORIGIN

```
Query Match          90.0%; Score 18; DB 9; Length 133984;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGGAAGGAGGAGGCGTCG 18
Db 31684 AGGAAGGAGGAGGCGTCG 31667
```

## RESULT 5

AC021462/c

LOCUS

```
DEFINITION Homo sapiens clone RP11-276C1, WORKING DRAFT SEQUENCE, 21 unordered
pieces.
```

ACCESSION

AC021462

VERSION

AC021462.3

KEYWORDS

HTG; HTGS PHASE3; HTGS DRAFT.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 153023)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens, clone RP11-276C1

## TITLE

JOURNAL

COMMENT

Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavsky, L., Boukigalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArillano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehotsky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Melidrom, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 3, 2000 this sequence version replaced GI:7230200.

All repeats were identified using RepeatMasker:

Smit, A. F. A. &amp; Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L5201

Center clone name: 276 C1

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 141815 bases at least Q40

Consensus quality: 147368 bases at least Q30

Consensus quality: 149481 bases at least Q20

Insert size: 157000; agarose-fp

Insert size: 151023; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1015: contig of 1015 bp in length
* 1015: gap of 100 bp
* 1116: contig of 1161 bp in length
* 2277: gap of 100 bp
* 2377: contig of 1633 bp in length
* 4010: gap of 100 bp
* 4110: contig of 1628 bp in length
* 5738: gap of 100 bp
* 5838: contig of 1102 bp in length
* 6940: gap of 100 bp
* 7040: contig of 2465 bp in length
* 9505: gap of 100 bp
* 9605: contig of 2130 bp in length
* 11735: gap of 100 bp
* 11835: contig of 2277 bp in length
* 14111: gap of 100 bp
* 14112: contig of 2762 bp in length
* 16974: gap of 100 bp
* 17074: contig of 2515 bp in length
* 19589: gap of 100 bp
* 19689: contig of 2587 bp in length
* 22276: gap of 100 bp
```

```

* 22376 25923: contig of 3448 bp in length
* 25924 25923: gap of 100 bp
* 25924 31307: contig of 5384 bp in length
* 31308 31407: gap of 100 bp
* 31408 34863: contig of 3456 bp in length
* 34864 34963: gap of 100 bp
* 34964 41382: contig of 6459 bp in length
* 41383 41482: gap of 100 bp
* 41483 49024: contig of 7542 bp in length
* 49025 49124: gap of 100 bp
* 49125 58672: contig of 9548 bp in length
* 58673 58772: gap of 100 bp
* 58773 69622: contig of 10850 bp in length
* 69623 69722: gap of 100 bp
* 69723 88191: contig of 18469 bp in length
* 88192 88291: gap of 100 bp
* 88292 107084: contig of 18793 bp in length
* 107085 107184: gap of 100 bp
* 107185 153023: contig of 45839 bp in length.

```

## FEATURES

source

```

1..153023
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-276C1"
/clone_lib="RPC1-11 Human Male BAC"

```

```

misc_feature 1..1015
              /note="assembly_fragment"
misc_feature 1116..2276
              /note="assembly_fragment"
misc_feature 2377..4009
              /note="assembly_fragment"
misc_feature 4110..5737
              /note="assembly_fragment"
misc_feature 5838..6939
              /note="assembly_fragment"
misc_feature 7040..9504
              /note="assembly_fragment"
misc_feature 9605..11734
              /note="assembly_fragment"
misc_feature 11835..14111
              /note="assembly_fragment"
misc_feature 14212..16973
              /note="assembly_fragment"
misc_feature 17074..19588
              /note="assembly_fragment"
misc_feature 19689..22275
              /note="assembly_fragment"
misc_feature 22376..25823
              /note="assembly_fragment"
misc_feature 25924..31307
              /note="assembly_fragment"
misc_feature 31408..34863
              /note="assembly_fragment"
misc_feature 34964..41382
              /note="assembly_fragment"
misc_feature 41483..49024
              /note="assembly_fragment"
misc_feature 49125..58672
              /note="assembly_fragment"
misc_feature 58773..69622
              /note="assembly_fragment"
misc_feature 69723..88191
              /note="assembly_fragment"
misc_feature 88292..107084
              /note="assembly_fragment"
misc_feature 107185..153023
              /note="assembly_fragment"
              clone_end:T7
              vector_side:left

```

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 153023;

```

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAAGAGAGAGCGCTCG 18
   |||||
DB 135271 AGGAAGAGAGAGCGCTCG 135254

RESULT 6
CER186 29134 bp DNA linear INV 10-DEC-2003
Caenorhabditis elegans cosmid R186, complete sequence.
ACCESSION Z78016
VERSION Z78016.1 GI:1483278
KEYWORDS HTG; Gephyrin like; Potassium channel protein; Signal recognition
          particle receptor beta subunit like.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
          Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1
AUTHORS none.
TITLE Genome sequence of the nematode C. elegans: a platform for
        investigating biology. The C. elegans Sequencing Consortium
        Science 282 (5396), 2012-2018 (1998)
JOURNAL 99069613
MEDLINE The C.elegans Sequencing Consortium.
REMARK 2 (bases 1 to 29134)
AUTHORS Barlow,K.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1996) Nematode Sequencing Project, Sanger
        Institute, Hinxton, Cambridge CB10 1SA, England and Department of
        Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
        jes@sanger.ac.uk or rw@nematode.wustl.edu
COMMENT Coding sequences below are predicted from computer analysis, using
        predictions from GeneFINDER (P. Green, U. Washington), and other
        available information.
        Current sequence finishing criteria for the C. elegans genome
        sequencing consortium are that all bases are either sequenced
        unambiguously on both strands, or on a single strand with both a
        dye primer and dye terminator reaction, from distinct subclones.
        Exceptions are indicated by an explicit note.
        For a graphical representation of this sequence and its analysis
        see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
        name=R186
        IMPORTANT: This sequence is NOT necessarily the entire insert of
        the specified clone. It may be shorter because we only sequence
        overlapping sections once, or longer because we arrange for a small
        overlap between neighbouring submissions.
        IMPORTANT: This sequence is not the entire insert of clone R186. It
        may be shorter because we only sequence overlapping sections once,
        or longer because we arrange for a small overlap between
        neighbouring submissions.
        The true left end of clone R186 is at 1 in this sequence. The true
        left end of clone T1593 is at 29032 in this sequence. The true
        right end of clone T1631 is at 1444 in this sequence. The start of
        this sequence (1..100) overlaps with the end of sequence Z81592.
        The end of this sequence (29032..29134) overlaps with the start of
        sequence Z78013.
FEATURES
        Location/Qualifiers
          1..29134
            /organism="Caenorhabditis elegans"
            /mol_type="genomic DNA"
            /strain="Bristol N2"
            /db_xref="taxon:6239"
            /chromosome="IV"
            /clone="R186"
            join(complement(101..558),
              complement(Z81592..34909..35209))
            /gene="R186.6"
            join(complement(101..558),
              complement(Z81592..34909..35209))
            /gene="R186.6"
          gene
          CDS

```

```

/standard_name="R186.6"
/notes="contains similarity to Pfam domain: PF00581
(Rhodanese-like domain), Score=42.7, E-value=2.7e-09, N=1"
/codon_start=1
/product="Hypothetical protein R186.6"
/db_xref="GI:7160720"
/translation="MESLLSKRDMENWTPNHLTLTLTKKICILEAASGDEAKSED
AFSDVHIESARLIHSHNLHAGVPVHPLOQRFARSQGINCHVIVYDRGQNTWSY
ITWIFKLFGRKVSLLGGYLGWKTHQARSQYITEQDAPRKPRQDGFELASMNDSVI
ITDVLINSEIDNFDVADTKDEFLGTAGALYGHIKGARNIPVDAVDWVGQWK
DADHLKGLFNNAISLRKPVVCYIVISQNFDPKS"
complement(join(2877..3009,3322..3577,3621..3700,
4050..4281,4342..4471))
/gene="R186.1"
complement(join(2877..3009,3322..3577,3621..3700,
4050..4281,4342..4471))
/gene="R186.1"
/standard_name="R186.1"
/notes="Weak similarity to Human T10 mRNA product
(TR.G406263)
cDNA EST Yk235e12.5 comes from this gene
cDNA EST Yk734f3.5 comes from this gene
cDNA EST Yk734f3.3 comes from this gene
cDNA EST Yk795b01.5 comes from this gene
cDNA EST Yk235e12.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein R186.1"
/protein_id="CAB01439.1"
/db_xref="GI:3879196"
/translation="MCITFTIKTAKSSLDKYKLIILNNRDEOLLPTKMEHWDGILSG
VDKQENARGTWLGNENGRIGMLSITQTOESXNLHAPSRGGIVNEFLNADTKM13
SLKVCASKGQFLAVKESKSTGLYEVRTLANOQVDIELVQLXDEHYVNSPTPK
YQKAVQKKLLREHLENSQDSVDQIEKLLSIKANTQWFDQLOLQIQONVEEYNR
PLSAIFIKYPGTMYGTRCHTLITVDQKINILRLRLPEQQTWHDARPEVLNGS
"
join(5527..5856,6303..6443,6489..6682,6764..6921,
6965..7152)
/gene="srd-35"
join(5527..5856,6303..6443,6489..6682,6764..6921,
6965..7152)
/gene="srd-35"
/standard_name="R186.2"
/codon_start=1
/product="C. elegans SRD-35 protein (corresponding
sequence R186.2)"
/protein_id="CAB01441.1"
/db_xref="GI:3879198"
/translation="MVPLQNVDEPSSKIYEVINEIVYPLITLIFHSVLVFIIF
FSPSHNLFPLTLTSLDVFATSIIFYIQPRLISEYQVNPVYCYGCPREKNTQFC
FVLPMMQTVSIAGISITVTLYPKYHLNLSKGLFGKLVRSILFLYIPCLSMCCA
FVVLNVLPNEELNKSFSDFTRTEFTKIGNMKLGELPNINFAMIAVALYCSLITF
WEFKTKLSLNALSGSSPVLYRHAQVMGKITIQVMHPIFYIPFTLYSYSLITGT
KIFVQQPFMAMSPNLSASFPLINLYFVVFIRIKSNFNGKHOQTSTIRIASLTPSG
YF"
complement(join(7346..7427,7474..7673,7776..7886,
7938..8071,8121..8223,8267..8359))
/gene="R186.3"
complement(join(7346..7427,7474..7673,7776..7886,
7938..8071,8121..8223,8267..8359))
/gene="R186.3"
/standard_name="R186.3"
/notes="Similarity to Mouse signal recognition particle
receptor beta subunit (SW:SRPB_MOUSE)
cDNA EST Yk522a4.3 comes from this gene
cDNA EST Yk522a4.5 comes from this gene
cDNA EST Yk807b04.3 comes from this gene
cDNA EST Yk807b04.5 comes from this gene
cDNA EST Yk1275a10.5 comes from this gene
cDNA EST Yk1275a10.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein R186.3"
/protein_id="CAB01443.1"

```

```

/db_xref="GI:3879200"
/translation="MDKIDFNDEFTTAVLATAVIGLITVLLLVLLKSPASSKNKRVLFV
GLMDCRKTITFOLSQKEASYPPTTKYTSWENKNTLRLKQKEKIIIDYFGNDRLRQ
KLJENHSLRSLRIIVFVDSAAFSKNARVAELFVALENVKVPILLACHKQDLIS
LAKTEKVRNSLEKEIGLINSRAAALIGTDSSEKRSITLDTGIDFKWEDLKKQEVFS
FVSTSGNSDFGVHEIASFVRA"
complement(join(8532..8626,8775..8878,8928..8975,
9541..9644))
/gene="R186.8"
complement(join(8532..8626,8775..8878,8928..8975,
9541..9644))
/gene="R186.8"
/standard_name="R186.8"
/codon_start=1
/product="Hypothetical protein R186.8"
/protein_id="CAB17911.1"
/db_xref="GI:33300311"
/translation="MARGIHLQHVFTKELVESTNNGLKCKILKRIYSFKLPLLYOK
VMGAKSYVIVQLASVITGSTRVWRERAKEFAGIPDPAYGKSCLFEEVKRVKGS
ELPKRIRGIYSIEN"
join(10042..10065,10127..10307,10591..10807,10916..11285,
11345..11728)
/gene="lin-46"
join(10042..10065,10127..10307,10591..10807,10916..11285,
11345..11728)
/gene="lin-46"
/standard_name="R186.4"
/notes="Similarity to Rat gephyrin (SW:GEPH_RAT), contains
similarity to Pfam domain: PF00994 (Molybdenum cofactor
biosynthesis protein), Score=82.9, E-value=4.4e-10, N=1
cDNA EST CEESK51FB comes from this gene
cDNA EST AF042069 comes from this gene
cDNA EST Yk1087f07.5 comes from this gene
cDNA EST Yk1087f07.3 comes from this gene"
/codon_start=1
/product="C. elegans LIN-46 protein (corresponding
sequence R186.4)"
/protein_id="CAB01440.2"
/db_xref="GI:14530531"
/translation="MSSGLKKEATLDDVFQKLEDLCKLPPOEKTVNTSLKTKGRIL
AEDITEYDIPQATISVDGFAIIIVQLGTRKRIIVGLSTAVTPYNAELISNECVRIIT
GGVVPDQATVPVEINVALEKKEKCEIVLKKPEGDNIREVSEAKTGEILLKQGHHL
DTMSITLHALGISQVEIYKKPRVCVLSIGSDLSNKMVGFNRSQILLELFSQGGTTA
IDAGSSHTITEVEEKIRTAASACVAVTVGGAAQVIREVAKTLKPKFIEDVDSTPGN
FTVSTGKIDETPVLSIFEPVHYVSMGANLFPVSPILRAMEGONSETSHRFAELTOP
ISKTSRTPRARSEVSKGNLISFPGCEDIPGANSILELVKSNCTCSAGVVDLRFPA"
join(12268..12365,12414..12566,12619..12982,13027..13258,
13928..14160)
/gene="R186.7"
join(12268..12365,12414..12566,12619..12982,13027..13258,
13928..14160)
/gene="R186.7"
/standard_name="R186.7"
/notes="cDNA EST Yk417a1.3 comes from this gene
cDNA EST Yk424g3.3 comes from this gene
cDNA EST Yk497g8.3 comes from this gene
cDNA EST Yk290a11.5 comes from this gene
cDNA EST Yk417a1.5 comes from this gene
cDNA EST Yk424g3.5 comes from this gene
cDNA EST Yk499a4.5 comes from this gene
cDNA EST Yk499a4.3 comes from this gene
cDNA EST Yk517d12.5 comes from this gene
cDNA EST Yk517d12.3 comes from this gene
cDNA EST Yk522f6.5 comes from this gene"

```

```

Query Match      87.0%; Score 17.4; DB 3; Length 29134;
Best Local Similarity 94.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AGAAAGGAGAGCGTCGT 19
D5      18183 ACGAAGGAGAGCGTCGT 18201

```

```

RESULT 7
AC020966/c
LOCUS      124479 bp      DNA      linear      HTG 10-FEB-2000
DEFINITION Mus musculus clone RP21-95D23, WORKING DRAFT SEQUENCE, 34 unordered
            pieces.
ACCESSION  AC020966
VERSION     1      GI:6691264
KEYWORDS    HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 124479)
            DOE Joint Genome Institute.
            Sequencing of Mouse
            Unpublished
            2 (bases 1 to 124479)
            DOE Joint Genome Institute.
            Direct Submission
            Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----Summary Statistics
            Consensus quality: 88445 bases at least Q40
            Consensus quality: 107058 bases at least Q30
            Consensus quality: 113897 bases at least Q20
            Estimated insert size: 124479; sum-of-contigs estimation
            Estimated insert size: 130000; pulse field gel estimation
            Quality coverage: 4.91x in Q20 bases; pulse field gel estimation
            Quality coverage: 5.13x in Q20 bases; sum-of-contigs estimation
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 34 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence.
            * as soon as it is available and the accession number will
            * be preserved.
            1
            1680: contig of 1680 bp in length
            gap of unknown length
            1681 2863: contig of 1183 bp in length
            gap of unknown length
            2864 4267: contig of 1404 bp in length
            gap of unknown length
            4268 5672: contig of 1405 bp in length
            gap of unknown length
            5673 7501: contig of 1829 bp in length
            gap of unknown length
            7502 8693: contig of 1192 bp in length
            gap of unknown length
            8694 10399: contig of 1706 bp in length
            gap of unknown length
            10400 11622: contig of 1223 bp in length
            gap of unknown length
            11623 12890: contig of 1268 bp in length
            gap of unknown length
            12891 15104: contig of 2214 bp in length
            gap of unknown length
            15105 16326: contig of 1222 bp in length
            gap of unknown length
            16327 18091: contig of 1765 bp in length
            gap of unknown length
            18092 19366: contig of 1275 bp in length
            gap of unknown length
            19367 20701: contig of 1335 bp in length
            gap of unknown length
            20702 22100: contig of 1399 bp in length
            gap of unknown length

```

```

* 22101 23518: contig of 1418 bp in length
* gap of unknown length
* 23519 26001: contig of 2483 bp in length
* gap of unknown length
* 26002 27358: contig of 1357 bp in length
* gap of unknown length
* 27359 29117: contig of 1759 bp in length
* gap of unknown length
* 29118 31049: contig of 1932 bp in length
* gap of unknown length
* 31050 34106: contig of 3057 bp in length
* gap of unknown length
* 34107 36494: contig of 2388 bp in length
* gap of unknown length
* 36495 39240: contig of 2746 bp in length
* gap of unknown length
* 39241 42051: contig of 2811 bp in length
* gap of unknown length
* 42052 43986: contig of 1935 bp in length
* gap of unknown length
* 43987 46160: contig of 2174 bp in length
* gap of unknown length
* 46161 50962: contig of 4802 bp in length
* gap of unknown length
* 50963 58010: contig of 7048 bp in length
* gap of unknown length
* 58011 64019: contig of 6009 bp in length
* gap of unknown length
* 64020 69890: contig of 5871 bp in length
* gap of unknown length
* 69891 77094: contig of 7204 bp in length
* gap of unknown length
* 77095 85512: contig of 8418 bp in length
* gap of unknown length
* 85513 103051: contig of 17539 bp in length
* gap of unknown length
* 103052 124479: contig of 21428 bp in length.
* Location/Qualifiers
* 1..124479
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /clone="RP21-95D23"
FEATURES             source
ORIGIN
Query Match      85.0%; Score 17; DB 2; Length 124479;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 2 GGAAAGGAGAGGCGTCG 18
    |||||
Db 15243 GGAAAGGAGAGGCGTCG 15227
RESULT 8
TTY14655
LOCUS      TTY14655      1014 bp      DNA      linear      BCT 23-APR-1998
DEFINITION Thermoproteus tenax Pfp gene.
ACCESSION  Y14655
VERSION    Y14655.1      GI:3087895
KEYWORDS   Pfp gene; pyrophosphate-dependent phosphofructokinase.
SOURCE     Thermoproteus tenax
            ORGANISM
            Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
            Thermoproteaceae; Thermoproteus.
            1
            Siebers, B., Klenk, H.P. and Hensel, R.
            Pfp-dependent phosphofructokinase from Thermoproteus tenax, an
            archaeal descendant of an ancient line in phosphofructokinase
            evolution
            J. Bacteriol. 180 (8), 2137-2143 (1998)
            98215182
            PUBMED 9555837

```

```

REFERENCE
AUTHORS      Siebers,B.
TITLE        Direct Submission
JOURNAL      Submitted (21-AUG-1997) B. Siebers, FB 9 Mikrobiologie,
              Universitaet-GH Essen, Universitaetsstrasse 5, 45117 Essen, FRG
FEATURES
  source
    1..1014
      /organism="Thermoproteus tenax"
      /mol_type="genomic DNA"
      /strain="Kral (DSM 2078)"
      /db_xref="taxon:2271"
    1..1014
      /gene="pfp"
      /gene="pfp"
    1..1014
      /gene="pfp"
      /EC_number="2.7.1.90"
      /note="pyrophosphate--fructose-6-phosphate
      1-phosphotransferase"
      /codon_start=1
      /transl_table=11
      /product="alternative product name:
      pyrophosphate-dependent phosphofructokinase"
      /protein_id="CAA74985.1"
      /db_xref="GI:3087896"
      /db_xref="GOA:O57694"
      /db_xref="SPTREMBL:O57694"
      /translation="MKIGVIGGGDAPGLNIATVTVKLAERKHEVVAIYHGWGLLN
      KEVKVSRDLDDFAFGGTVIRTSRTNPFKDEERARLLENVRELGLDVVAATGGDD
      TLGAAGARQGIIDAVGIPKTIIDNVYGTDTTIGFDSAVNAAIEATESPFTLISHE
      RIGVYVNGRAGAIATLFTGLSTWADAVLIPEFASWDSVAKRVEAYNERMALVVV
      SEGIYGGPKDEYGHGFLGVGNELAEVIIRSTGIBARAVVLGHTIRGVPPTAFDR
      LAVRYATAAYEAVENTGRYGVNVAISNGDIATVPIVDVVGKNLVSGYWMRLIYETWPD
      LAG"
ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 1014;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAAGGAGAGCGTCGTC 20
    |||||
Db 476 ACGAAAGCATAGGCGTCGTC 495

RESULT 9
CHKPLB/c      3312 bp mRNA linear VRT 06-MAR-1995
LOCUS         Chicken cardiac phospholamban (plb) mRNA, complete cds.
DEFINITION
ACCESSION     M59039
VERSION       M59039.1 GI:212575
KEYWORDS      phospholamban.
SOURCE        Gallus gallus (chicken)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE     1 (bases 1 to 3312)
AUTHORS       Toyofuku,T. and Zak,R.
TITLE        Characterization of cDNA and genomic sequences encoding a chicken
              phospholamban
JOURNAL       J. Biol. Chem. 266 (9), 5375-5383 (1991)
MEDLINE       91170195
PUBMED        1825996
COMMENT       Original source text: Chicken (broiler breeders) 7-week old adult
              heart, cDNA to mRNA, clones CPL-[6,12,15].
FEATURES
  source
    1..3312
      /organism="Gallus gallus"
      /mol_type="mRNA"
      /strain="broiler breeder"
      /db_xref="taxon:9031"
      /clone="CPL-15"
      /clone="CPL-6"

/clone="CPL-12"
/tissue_type="heart"
/dev_stage="7-week old adult"
/tissue_lib="ZAP cDNA library"
1..3312
/gene="plb"
/gene="plb"
246..404
/gene="plb"
/codon_start=1
/product="phospholamban"
/protein_id="AAA62738.1"
/db_xref="GI:212576"
/translation="MEKQVYITRSALRRASTLEVNPQARQRLQELFVNFCLILCLLL
ICLIIVMLL"
571..576
/gene="plb"
604
/gene="plb"
966..971
/gene="plb"
987
/gene="plb"
3071..3312
/gene="plb"
/note="putative VECTOR sequence Bacteriophage lambda
(J02459); putative"
3290..3295
/gene="plb"
3312
/gene="plb"
ORIGIN
Query Match      84.0%; Score 16.8; DB 5; Length 3312;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAAGGAGAGCGTCGTC 20
    |||||
Db 202 ACGAAAGGAGAGGAGTGTGTC 183

RESULT 10
CHKPLB2/c      3423 bp DNA linear VRT 07-MAR-1995
LOCUS         Chicken phospholamban (plb) gene, exon 2.
DEFINITION
ACCESSION     M59038
VERSION       M59038.1 GI:212578
KEYWORDS      phospholamban.
SEGMENT       2 of 2
SOURCE        Gallus gallus (chicken)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE     1 (bases 1 to 3423)
AUTHORS       Toyofuku,T. and Zak,R.
TITLE        Characterization of cDNA and genomic sequences encoding a chicken
              phospholamban
JOURNAL       J. Biol. Chem. 266 (9), 5375-5383 (1991)
MEDLINE       91170195
PUBMED        1825996
COMMENT       Original source text: Chicken (White leghorn) adult liver DNA,
              clones GPL-1. and GPL-2.
FEATURES
  source
    1..3423
      /organism="Gallus gallus"
      /mol_type="genomic DNA"
      /strain="White leghorn"
      /sub_species="domesticus"
      /db_xref="taxon:9031"
      /clone="GPL-1"
      /tissue_type="liver"
      /dev_stage="adult"

```

```

/tissue lib="PEMBL3 chicken genomic"
join(M59037.1:511..1803,1..3423)
/gene="plb"
mRNA
join(M59037.1:1621..1745,89..3423)
/gene="plb"
/product="phospholamban"
join(M59037.1:1621..1745,89..3276)
/gene="plb"
/product="phospholamban"
join(M59037.1:1621..1745,89..951)
/gene="plb"
/product="phospholamban"
join(M59037.1:1621..1745,89..567)
/gene="plb"
/product="phospholamban"
order(M59037.1:1746..1803,1..88)
/gene="plb"
89..3423
/gene="plb"
/number=2
89..3276
/gene="plb"
/number=2
89..951
/gene="plb"
/number=2
89..567
/gene="plb"
/number=2
209..367
/gene="plb"
/codon_start=1
/product="phospholamban"
/protein_id="AAA63167.1"
/db_xref="GI:212580"
/translation="MEKVVQYITRSLRASTLEVNPPQRLQELFVNFCLILICLL
ICILVLL"
534..539
/gene="plb"
930..935
/gene="plb"
3254..3259
/gene="plb"

```

```

ORIGIN
Query Match      84.0%; Score 16.8; DB 5; Length 3423;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAGAGAGAGCGTCCTC 20
|||||
Db 165 AGGAAGAGAGAGAGTTC 146

RESULT 11
AK065803
LOCUS AK065803
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013036D01, full
insert sequence.
ACCESSION AK065803
VERSION AK065803.1 GI:32975821
KEYWORDS Full_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, I., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

```

```

Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Otono, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Nikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Komoto, K., Miyazaki, A., Oota, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 3492)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Kagawa, I.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kaga, J.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Oka, M., Ooka, H.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Sakai, K.,
Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N.,
Oka, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 128.353 Seconds  
(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819b-35

Perfect score: 20

Sequence: 1 aggaagagagagcgctgc 20

Scoring table: IDENTITY\_NUC

Gapop 1.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : N Geneseq 29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1980s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003as:\*

8: geneseqn2003bs:\*

9: geneseqn2003cs:\*

10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	20	100.0	20	6	AAL48167 Human HLA
2	20	100.0	20	6	AAL48199 Human HLA
3	20	100.0	20	6	AAL48197 Human HLA
4	20	100.0	20	6	AAL48183 Human HLA
5	16.8	84.0	48436	6	ABN89533 Human cor
6	16.8	84.0	160552	4	AAD02697 Human gly
7	16.4	82.0	546	4	ABL22959 Drosophila
8	16.4	82.0	2546	4	ABL22958 Drosophila
9	15.8	79.0	307	6	ABL98955 Omega-con
10	15.8	79.0	850	6	ABQ49803 Oligonucle
11	15.8	79.0	850	6	ABQ49802 Oligonucle
12	15.8	79.0	1362	6	ABK72943 Bacillus
13	15.8	79.0	1686	5	AS85172 DNA encod
14	15.8	79.0	2000	7	ADA73217 Rice gene
15	15.8	79.0	2000	7	ADA72949 Rice gene
16	15.8	79.0	3213	7	ABZ24255 Human SLC
17	15.8	79.0	3213	9	ADD18763 Human dis
18	15.8	79.0	3229	5	AS75476 DNA encod
19	15.8	79.0	3229	5	AS75478 DNA encod
20	15.8	79.0	6901	9	ADE09927 Novel DNA
21	15.8	79.0	7690	6	ABL31122 Human imm
22	15.8	79.0	14924	6	ABL32225 Human imm
23	15.8	79.0	14924	6	ABL54322 Chemicall

24	15.8	79.0	177851	7	AAL57272	AAL57272 ba438B23-
25	15.4	77.0	296	8	ACL21232	ACL21232 DNA clone
26	15.4	77.0	525	8	ACL21232	ACL21232 DNA clone
27	15.4	77.0	558	8	ACL21234	ACL21234 DNA clone
28	15.4	77.0	562	8	ACL21236	ACL21236 DNA clone
29	15.4	77.0	574	8	ACL21228	ACL21228 DNA clone
30	15.4	77.0	610	6	ABQ15884	ABQ15884 Oligonucle
31	15.4	77.0	610	6	ABQ15885	ABQ15885 Oligonucle
32	15.4	77.0	673	6	ABQ42042	ABQ42042 Oligonucle
33	15.4	77.0	673	6	ABQ42043	ABQ42043 Oligonucle
34	15.4	77.0	713	8	ACL21235	ACL21235 DNA clone
35	15.4	77.0	716	8	ACL21233	ACL21233 DNA clone
36	15.4	77.0	852	6	ABN68486	ABN68486 Streptoco
37	15.4	77.0	1317	2	AA59113	AA59113 Mouse pan
38	15.4	77.0	1522	2	AA59115	AA59115 Mouse pan
39	15.4	77.0	2016	9	ADB90631	ADB90631 Guanine n
40	15.4	77.0	2402	9	ADB90630	ADB90630 Guanine n
41	15.4	77.0	5524	9	ADB90625	ADB90625 Mouse gua
42	15.4	77.0	34118	9	ADC86456	ADC86456 Human GPC
43	15.4	77.0	110000	7	AAL52246	Continuation (2 of
44	15.4	77.0	227968	6	ABK83497	ABK83497 Human CDN
45	15.2	76.0	65	6	ABN27551	ABN27551 Rat splic
46	15.2	76.0	323	5	ABV48086	ABV48086 Human pro
47	15.2	76.0	459	8	ACH38985	ACH38985 Human foe
48	15.2	76.0	461	5	ABV18299	ABV18299 Human pro
49	15.2	76.0	488	8	ACH13744	ACH13744 Human adu
50	15.2	76.0	590	3	AAFI4812	AAFI4812 Aspergill
51	15.2	76.0	590	7	ABZ54025	ABZ54025 Aspergill
52	15.2	76.0	776	6	ABQ48378	ABQ48378 Oligonucle
53	15.2	76.0	776	6	ABQ48379	ABQ48379 Oligonucle
54	15.2	76.0	780	6	ABQ49844	ABQ49844 Oligonucle
55	15.2	76.0	780	6	ABQ49845	ABQ49845 Oligonucle
56	15.2	76.0	949	4	AAH05625	AAH05625 Human CDN
57	15.2	76.0	1156	6	ABK63440	ABK63440 Rat seque
58	15.2	76.0	1156	9	ADB57852	ADB57852 Toxicity-
59	15.2	76.0	1156	9	ADB52358	ADB52358 Primary r
60	15.2	76.0	1156	9	ADB85151	ADB85151 Rat UDP-g
61	15.2	76.0	1212	6	ABQ69043	ABQ69043 Listeria
62	15.2	76.0	1453	3	AA42323	AA42323 Arabidops
63	15.2	76.0	1584	4	AAH14112	AAH14112 Human CDN
64	15.2	76.0	1916	4	AA512684	AA512684 Human CDN
65	15.2	76.0	1937	4	AA51307	AA51307 Human pol
66	15.2	76.0	1938	4	ABA08798	ABA08798 Human cor
67	15.2	76.0	2505	7	ABZ23069	ABZ23069 Human GPC
68	15.2	76.0	2700	4	AAH18502	AAH18502 Human CDN
69	15.2	76.0	2701	5	ABAL5977	ABAL5977 Human ner
70	15.2	76.0	2934	4	ABL15887	ABL15887 Drosophila
71	15.2	76.0	3396	2	AAQ55515	AAQ55515 PTPase PT
72	15.2	76.0	3561	2	AAQ55970	AAQ55970 PTPase PT
73	15.2	76.0	3566	6	ABI92337	ABI92337 Mouse isc
74	15.2	76.0	4608	6	ABQ70898	ABQ70898 Listeria
75	15.2	76.0	4841	4	AAK52955	AAK52955 Human pol
76	15.2	76.0	4880	4	AAK51971	AAK51971 Human pol
77	15.2	76.0	4898	7	ABZ79896	ABZ79896 Human nuc
78	15.2	76.0	5204	4	ABL15886	ABL15886 Drosophila
79	15.2	76.0	12961	5	ABA20012	ABA20012 Human ner
80	15.2	76.0	18890	4	ABL13818	ABL13818 Drosophila
81	15.2	76.0	24606	4	AAK84723	AAK84723 Human imm
82	15.2	76.0	24606	4	AAK78367	AAK78367 Human imm
83	15.2	76.0	28564	9	ADE63609	ADE63609 Human gen
84	15.2	76.0	39887	4	AAK81263	AAK81263 Human imm
85	15.2	76.0	39887	4	AAK79153	AAK79153 Human imm
86	15.2	76.0	70419	7	AAK56111	AAK56111 Human WNT
87	15.2	76.0	70419	8	ADA02473	ADA02473 Human WNT
88	15.2	76.0	70419	9	ADB72212	ADB72212 Human WNT
89	15.2	76.0	96583	3	AAF22297	AAF22297 BAC conta
90	15.2	76.0	110000	6	ABX08336	ABX08336 07
91	15.2	76.0	110000	6	ABQ67196	ABQ67196 5
92	15.2	76.0	110000	6	ABQ69245	ABQ69245 26
93	15.2	76.0	110000	6	ABA03041	ABA03041 26
94	15	75.0	2063	9	ADB63557	ADB63557 Human CDN
95	15	75.0	2970	9	ACF79497	ACF79497 Cattle pl
96	15	75.0	249487	6	ABN85733	ABN85733 Mouse gen



c 97 14.8 74.0 235 4 AAH36766 Human col  
 c 98 14.8 74.0 271 7 ABX84599 Corn ear-  
 c 99 14.8 74.0 284 7 ABX88350 Corn ear-  
 c 100 14.8 74.0 345 7 ACF68272  
 c 101 14.8 74.0 524 6 ACF79924 Bacillus  
 c 102 14.8 74.0 525 7 ACF68664  
 c 103 14.8 74.0 564 9 ADD17229 DNA (Seq)  
 c 104 14.8 74.0 564 9 ADD16331 DNA (Seq)  
 c 105 14.8 74.0 597 5 AAS71493  
 c 106 14.8 74.0 621 6 ABX73246 Bacillus  
 c 107 14.8 74.0 642 6 ABQ17328 Oligonucle  
 c 108 14.8 74.0 642 6 ABQ17329 Oligonucle  
 c 109 14.8 74.0 732 2 AAQ23090 Antigen a  
 c 110 14.8 74.0 739 4 AAH07144 Human cDN  
 c 111 14.8 74.0 801 6 ABQ42574 Oligonucle  
 c 112 14.8 74.0 801 6 ABQ42575 Oligonucle  
 c 113 14.8 74.0 890 4 AAK61884 Human imm  
 c 114 14.8 74.0 911 6 ABQ42462 Oligonucle  
 c 115 14.8 74.0 911 6 ABQ42463 Oligonucle  
 c 116 14.8 74.0 1380 6 ABQ99350 Human cod  
 c 117 14.8 74.0 1754 7 ABT21008  
 c 118 14.8 74.0 1884 7 ACA31074  
 c 119 14.8 74.0 1937 7 ACC46489 Human dit  
 c 120 14.8 74.0 2030 7 ADA73451  
 c 121 14.8 74.0 2039 5 AAS81318  
 c 122 14.8 74.0 2318 4 AAH16372 Human cDN  
 c 123 14.8 74.0 2803 7 ACC79078 Human sec  
 c 124 14.8 74.0 2821 9 ADB62242 Human cDN  
 c 125 14.8 74.0 2966 3 AAA46689 DNA encod  
 c 126 14.8 74.0 3361 4 AAK81871 Human imm  
 c 127 14.8 74.0 3479 5 AAH21451  
 c 128 14.8 74.0 3480 3 AAA07601 Long QT s  
 c 129 14.8 74.0 3936 4 AAL07119 Human rep  
 c 130 14.8 74.0 3950 3 AAA07602 Long QT s  
 c 131 14.8 74.0 4070 4 AAF61965 Human eag  
 c 132 14.8 74.0 4070 4 AAL66257 Human poc  
 c 133 14.8 74.0 4070 7 ABZ76227 Human HER  
 c 134 14.8 74.0 4070 9 ADD02768 Human HER  
 c 135 14.8 74.0 4380 6 ABN83967 Human Gen  
 c 136 14.8 74.0 4710 4 ABL44084 Drosophil  
 c 137 14.8 74.0 5243 6 ABK52771 CDNA encod  
 c 138 14.8 74.0 5243 7 AAL53919 DNA of a  
 c 139 14.8 74.0 5401 4 AAL07117 Human rep  
 c 140 14.8 74.0 6169 4 AAK89619 Human dig  
 c 141 14.8 74.0 7028 4 ABL30122 Drosophil  
 c 142 14.8 74.0 7694 4 AAL03544 Human rep  
 c 143 14.8 74.0 7694 4 ABA07755 Human ova  
 c 144 14.8 74.0 8088 3 AAZ45540 Complete  
 c 145 14.8 74.0 8535 2 AAQ73731 Galv SEAT  
 c 146 14.8 74.0 9160 4 ABL20152 Drosophil  
 c 147 14.8 74.0 10069 6 ABK88427 Mouse T c  
 c 148 14.8 74.0 10213 9 ADD02779 AAK89617 Human dig  
 c 149 14.8 74.0 13310 4 AAK89617 Human dig  
 c 150 14.8 74.0 13310 4 AAK75442 Human imm  
 c 151 14.8 74.0 17481 4 AAK89618 Human imm  
 c 152 14.8 74.0 17481 4 AAK73255 Human dig  
 c 153 14.8 74.0 17481 4 AAK75443 Human imm  
 c 154 14.8 74.0 17481 4 AAK68583 Human imm  
 c 155 14.8 74.0 33239 7 ABX13675 Human sec  
 c 156 14.8 74.0 42334 8 ADA02588 Human ICS  
 c 157 14.8 74.0 42334 9 ADB72326  
 c 158 14.8 74.0 89328 6 ABL61995  
 c 159 14.8 74.0 110000 7 ACF67367 09  
 c 160 14.8 74.0 110000 7 ACF67367 14  
 c 161 14.8 74.0 110000 7 ACF65384 3  
 c 162 14.8 74.0 110000 7 ABL16390 0  
 c 163 14.8 74.0 249878 7 ACF65381  
 c 164 14.8 74.0 302250 6 ABL67703  
 c 165 14.8 74.0 349980 6 ABQ81847 Bifidobac  
 c 166 14.8 74.0 349980 6 ABQ81848 Bifidobac  
 c 167 14.4 72.0 330 6 ABN93034 Staphyloc  
 c 168 14.4 72.0 405 4 AAL86637 Human ova  
 c 169 14.4 72.0 412 6 ABL77754 Human poa

## ALIGNMENTS

RESULT 1  
 AAL48167  
 ID AAL48167 standard; DNA; 20 BP.  
 XX  
 AC AAL48167;  
 DT 01-OCT-2002 (first entry)  
 DE Human HLA DPB1 locus polymorphism address tag sequence #3.  
 XX  
 KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
 KW flow cytometry; human; DPB1; address tag; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200194639-A1.  
 PD 13-DEC-2001.  
 XX  
 PF 07-JUN-2001; 2001WO-US018590.  
 PR 08-JUN-2000; 2000US-0210759P.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX White PS, Torney DC;  
 PI WPI; 2002-566450/60.  
 DR  
 XX  
 PT Identifying sequences useful as address/capture tags for flow cytometry  
 PT based minisequencing, by generating tag sequences and rejecting sequences  
 PT based on certain parameters e.g. sequences which form stable hairpins.  
 XX Disclosure; Page 9; 35pp; English.

The present invention relates to a method of identifying sequences useful  
 as address/capture tags, involving rejecting sequences having common sub-  
 bases, and sequences which can form stable hairpins and stable dimers  
 from a sample of oligonucleotides, and selecting those sequences in the  
 sample that would hybridize to their respective complements with a high  
 degree of specificity. The method is useful for identifying a set of  
 sequences useful as address/capture tags which can be used for  
 multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
 cytometry assay. The present sequence is an address tag described in the  
 exemplification of the invention

Sequence 20 BP; 6 A; 3 C; 9 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;

Best Local Similarity 100.0%; Fred. No. 6.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAAAGGAGAGCGTCGTC 20

|||||

Db 1 AGGAAAGGAGAGCGTCGTC 20



PT Identifying sequences useful as address/capture tags for flow cytometry  
PT based minisequencing, by generating tag sequences and rejecting sequences  
PT based on certain parameters e.g. sequences which form stable hairpins.

XX Disclosure; Page 9; 35pp; English.

XX The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is an address tag described in the  
CC exemplification of the invention

XX Sequence 20 BP; 2 A; 9 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAAGGAGGAGGCGTCGTC 20  
Db 20 AGGAAGGAGGAGGCGTCGTC 1

RESULT 5

ABN89533/c  
ID ABN89533 standard; DNA; 48436 BP.

AC ABN89533;

DT 05-SEP-2002 (first entry)

DE Human corneal N-acetylglucosamine-6-sulfotransferase DNA SEQ ID NO:38.

KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;  
KW corneal; sulfatase; keratan sulfate; macular corneal dystrophy; MCD;  
KW ophthalmological; chromosome 16q22; gene; ds.

OS Homo sapiens.

PN US2002061562-A1.

XX 23-MAY-2002.

XX 09-AUG-2001; 2001US-00927602.

XX 11-AUG-2000; 2000US-00638211.

XX 11-AUG-2000; 2000US-0325773P.

XX (FUKU)/ FUKUDA M N.  
XX (AKAN)/ AKANA T O.

XX Fukuda MN, Akama TO;

XX WPI; 2002-507643/54.

XX New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase,  
PT useful for treatment, monitoring and diagnosis of macular corneal  
PT dystrophy.

XX Claim 35; Page 31-53; 69pp; English.

XX The present sequence encodes human corneal N-acetylglucosamine-6-  
CC sulfotransferase (GlcNAc6ST) (I), which is able to catalyse sulfation of  
CC keratan sulfate (KS). Also described is a method for monitoring the  
CC effect of treatments for macular corneal dystrophy (MCD), and detecting  
CC susceptibility to MCD. (I) is located to chromosome 16q22, and has  
CC ophthalmological activity. (I) can be used to treat or prevent macular

CC corneal dystrophy types I or II. (I) makes possible treatment of MCD  
CC without requiring keratoplasty or keratectomy

XX Sequence 48436 BP; 11653 A; 11904 C; 11645 G; 13234 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 6; Length 48436;

Best Local Similarity 90.0%; Pred. No. 2.8e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAGGAGGAGGCGTCGTC 20

Db 2137 AGGAAGGAGGAGGCGTCATC 2118

RESULT 6

AAD02697/c

ID AAD02697 standard; DNA; 160552 BP.

XX AAD02697;

XX 02-MAY-2001 (first entry)

XX Human glycosyl sulfotransferase-4 (GST-4) genomic DNA.

XX Human; glycosyl sulfotransferase-4; GST-4; immunosuppressive; therapy;  
KW selectin binding inhibitor; gene therapy; inflammation;  
KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;  
KW polyarthritis nodosa; polymyositis; systemic sclerosis; dermatitis;  
KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adenailitis;  
KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;  
KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;  
KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;  
KW asthma; hypersensitivity; rheumatic fever; tissue rejection;  
KW chromosome 16q23.1; ds.

XX Homo sapiens.

PH Key Location/Qualifiers

FT exon 32847..32922

FT /tag= a

FT /number= 1

FT /label= 4a SU4

FT 32923..35592

FT /tag= b

FT /cons\_splice= (5'site:NO, 3'site:YES)

FT 35593..35674

FT /tag= c

FT /number= 2

FT /label= 4a SU3

FT 35675..45093

FT /tag= d

FT 45094..45185

FT /tag= e

FT /number= 3

FT /label= 4a SU2

FT 45186..45633

FT /tag= f

FT /cons\_splice= (5'site:NO, 3'site:NO)

FT 45634..46700

FT /tag= g

FT /number= 4

FT /label= 4a SU1

FT 46701..47938

FT /tag= h

FT /cons\_splice= (5'site:YES, 3'site:NO)

FT 47939..49746

FT /tag= i

FT /number= 5

FT /note= "Includes 17 base pairs of 5'UTR, the ORF and all  
FT of 3'UTR"

FT 47939..47955

FT /tag= j

FT /note= "Portion of 5' untranslated region (5'UTR)"

CDS 47956..49128  
/\*tag= k  
/product= "Human glycosyl transferase-4alpha (GST-4alpha)"  
3'UTR 49129..49746  
/\*tag= l  
exon 83257..83347  
/\*tag= m  
/label= 4a\_502  
intron 83348..96412  
/\*tag= n  
/cons\_splice= (5'site:NO, 3'site:NO)  
exon 96413..96484  
/\*tag= o  
/label= 4a\_501  
intron 96485..98456  
/\*tag= p  
/cons\_splice= (5'site:NO, 3'site:NO)  
exon 98457..99968  
/\*tag= q  
/note= "Includes 17 base pairs of 5'UTR, the ORF and all of 3'UTR"  
5'UTR 98457..98473  
/\*tag= r  
/note= "Portion of 5' untranslated region. (5'UTR)"  
CDS 98474..99661  
/\*tag= s  
/product= "Human glycosyl transferase-4beta (GST-4beta)"  
3'UTR 99662..99968  
/\*tag= t  
WO200106015-A1.  
25-JAN-2001.  
19-JUL-2000; 2000WO-US019741.  
20-JUL-1999; 99US-01446942P.  
13-JUN-2000; 2000US-00593828.  
(REGC ) UNIV CALIFORNIA.  
Rosen SD, Lee JK, Hammerich S;  
WPI; 2001-138471/14.  
P-PSDE; AAY72639, AAY72640.  
New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications.  
Example 1; Page 62-104; 128pp; English.  
The present sequence is human glycosyl sulfotransferase-4 (GST-4) genomic DNA encoding GST-4alpha and GST-4beta. GST-4 gene is found on chromosome 16p23.1. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarthritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dermatitis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation  
SQ Sequence 160552 BP; 40281 A; 37573 C; 38015 G; 44564 T; 0 U; 119 Other;  
Query Match 84.0%; Score 16.8; DB 4; Length 160552;

Best Local Similarity 90.0%; Pred. No. 2.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGGAAAGCAGAGCGTCGTC 20  
||||| |||||||  
DB 53046 AGGAAAGCAGAGCGTCATC 53027  
RESULT 7  
ABL22959  
ID ABL22959 standard; DNA; 546 BP.  
XX  
AC ABL22959;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20350.  
XX  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
KW Drosophila melanogaster.  
OS  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers BW;  
XX WPI; 2001-656860/75.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.  
XX  
PS Claim 1; SEQ ID NO 20350; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
XX  
SQ Sequence 546 BP; 113 A; 193 C; 167 G; 73 T; 0 U; 0 Other;  
Query Match 82.0%; Score 16.4; DB 4; Length 546;  
Best Local Similarity 94.4%; Pred. No. 3.6e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 GAAAGGAGAGCGTCGTC 20  
||||| |||||||  
DB 449 GCAAGGAGAGCGTCGTC 466  
RESULT 8  
ABL22958  
ID ABL22958 standard; DNA; 2546 BP.  
XX  
AC ABL22958;  
XX  
DT 26-MAR-2002 (first entry)

```
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 20347.
XX DE
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PN 27-SEP-2001.
XX PD
XX PF 23-MAR-2001; 2001WO-US009231.
XX PF 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PR
XX XX (PEKE ) PE CORP NY.
XX FA
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI
XX XX WPI; 2001-656860/75.
XX DR
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 20347; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB161176-AB163851), expressed DNA
XX CC sequences (AB161840-AB161875) and the encoded proteins (ABBS7737-
XX CC ABBS72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX CC
XX SQ Sequence 2546 BP; 753 A; 609 C; 568 G; 616 T; 0 U; 0 Other;
XX
XX Query Match 82.0%; Score 16.4; DB 4; Length 2546;
XX Best Local Similarity 94.4%; Pred. No. 3.8e+02;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 3 GAAGGAGAGGCGTCCTC 20
XX Db 1449 GCAAGGAGAGGCGTCGTC 1466
XX
XX RESULT 9
XX ABL98955
XX ID ABL98955 standard; DNA; 307 BP.
XX AC ABL98955;
XX XX
XX DT 12-JUL-2002 (first entry)
XX XX
XX DE Omega-conopeptide Ra6.4 encoding DNA.
XX XX
XX KW Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;
XX KW neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;
XX KW antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;
XX KW anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;
XX KW neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;
XX KW stroke; cerebrovascular accident; brain trauma; spinal cord trauma;
XX KW drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;
XX KW migraine; inflammation; cardiovascular disorder; psychiatric disorder;
XX KW psychosis; anxiety; schizophrenia; ds.
XX OS
XX OS Conus rattus.
XX XX
```

```
PN WO200207675-A2.
XX XX
XX ED 31-JAN-2002.
XX XX
XX PF 23-JUL-2001; 2001WO-US023041.
XX XX
XX PF 21-JUL-2000; 2000US-0219616P.
XX PR 05-FEB-2001; 2001US-0265888P.
XX PR
XX XX (UTAH ) UNIV UTAH RES FOUND.
XX FA (COGN-) COGNETIX INC.
XX FA
XX PI Olivera BV, McIntosh JM, Watkins M, Garrett JE, Shon K;
XX PI Jacobsen R, Jones RM, Cartier GE;
XX PI
XX XX WPI; 2002-257318/30.
XX DR P-PSDE; ABB96696.
XX DR
XX PT New omega-conopeptides useful for treating disorders associated with
XX PT voltage gated ion channels e.g. pain, inflammation, neurologic or
XX PT cardiovascular disorders.
XX PS Claim 3; Page 70; 195pp; English.
XX CC The invention relates to isolated omega-conopeptides, nucleic acid
XX CC sequences encoding them, and propeptide sequences. The activity of the
XX CC peptides of the invention may be described as, analgesic, anticonvulsant,
XX CC vasotropic, cardiant, neuroprotective, cerebroprotective, cardiovascular,
XX CC antiinflammatory, antimigraine, antidiabetic, tranquiliser, vulnerary,
XX CC anxiolytic, neuroleptic, antipsychotic. Peptides of the invention act
XX CC by modulating the activity of voltage gated ion channels. They may be
XX CC used for treating or preventing disorders associated with voltage gated
XX CC ion channels such as neurological disorders, e.g. seizure (associated
XX CC with epilepsy), neurotoxic injury associated with conditions of hypoxia,
XX CC anoxia, ischaemia, stroke, cerebrovascular accident, brain or spinal
XX CC cord trauma, drowning, suffocation, perinatal asphyxia or hypoglycaemic
XX CC events; pain e.g. migraine; inflammation or cardiovascular disorders.
XX CC They may also be used for treating psychiatric disorders e.g. psychosis,
XX CC anxiety or schizophrenia. The analgesic agents of the invention show
XX CC diminished side effects and toxicity, and are non-addictive. The
XX CC sequences given in records ABL98855-ABL98956 represent omega-conopeptide
XX CC encoding DNA's
XX SQ Sequence 307 BP; 58 A; 84 C; 82 G; 83 T; 0 U; 0 Other;
XX
XX Query Match 79.0%; Score 15.8; DB 6; Length 307;
XX Best Local Similarity 89.5%; Pred. No. 6.7e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AGGAAGAGGAGGCGTCGT 19
XX Db 87 AGGTAAGGAGAGGCGTCGT 105
XX
XX RESULT 10
XX ABL9803/c
XX ID ABL9803 standard; DNA; 850 BP.
XX AC ABL9803;
XX XX
XX DT 12-JUL-2002 (first entry)
XX XX
XX XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 36394.
XX DE
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX XX
XX OS
XX OS Homo sapiens.
XX XX
XX PN WO200216632-A2.
XX XX
```

```

PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EP010074.
XX
XX 01-SEP-2000; 2003DE-01C43826.
XX
XX 05-SEP-2000; 2003DE-01C44543.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one member,
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
XX degree of hybridisation to both classes is determined from the label on
XX the amplicon. From the ratio of labels hybridised to the two classes of
XX oligomers, the degree of methylation is calculated. The method is used:
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
XX and of a wide range of diseases, e.g. cancer, disorders of the central
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
XX particularly by detecting mutations or single nucleotide polymorphisms
XX (SNP's); and (ii) for differentiation of cell or tissue types and for
XX investigating cell differentiation. The method allows the methylation
XX status of many C residues to be determined simultaneously. ABQ13410-
XX ACQ54121 represent genomic DNA sequences used to illustrate the method
XX for determining the degree of cytosine methylation described in the
XX disclosure of the invention
XX
XX Sequence 850 BP; 277 A; 340 C; 120 G; 113 T; 0 U; 0 Other;
XX
XX Query Match 79.0%; Score 15.8; DB 6; Length 850;
XX Best Local Similarity 89.5%; Pred. No. 7e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 GGAAAGGAGAGCGCTCGTC 20
XX ||| | ||||| |||||
XX Db 433 GGAGACGAGAGCGCTCGTC 415
XX
XX RESULT 11
XX ABQ49802
XX ID ABQ49802 standard; DNA; 850 BP.
XX
XX AC ABQ49802;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 36193.
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX WO200218632-A2.
XX
XX 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EP010074.

```

```

XX 01-SEP-2000; 2000DE-01C43826.
XX
XX 05-SEP-2000; 2000DE-01C44543.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one member,
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
XX degree of hybridisation to both classes is determined from the label on
XX the amplicon. From the ratio of labels hybridised to the two classes of
XX oligomers, the degree of methylation is calculated. The method is used:
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
XX and of a wide range of diseases, e.g. cancer, disorders of the central
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
XX particularly by detecting mutations or single nucleotide polymorphisms
XX (SNP's); and (ii) for differentiation of cell or tissue types and for
XX investigating cell differentiation. The method allows the methylation
XX status of many C residues to be determined simultaneously. ABQ13410-
XX ACQ54121 represent genomic DNA sequences used to illustrate the method
XX for determining the degree of cytosine methylation described in the
XX disclosure of the invention
XX
XX Sequence 850 BP; 113 A; 120 C; 340 G; 277 T; 0 U; 0 Other;
XX
XX Query Match 79.0%; Score 15.8; DB 6; Length 850;
XX Best Local Similarity 89.5%; Pred. No. 7e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 GGAAAGGAGAGCGCTCGTC 20
XX ||| | ||||| |||||
XX Db 418 GGAGACGAGAGCGCTCGTC 436
XX
XX RESULT 12
XX ABK72943
XX ID ABK72943 standard; DNA; 1362 BP.
XX
XX AC ABK72943;
XX
XX 13-AUG-2002 (first entry)
XX
XX Bacillus licheniformis genomic sequence tag (GST) #234.
XX
XX Differential gene expression; genomic sequenced tag; GST;
XX altered culture condition; environmental stress;
XX physiological provocation; ds.
XX
XX Bacillus licheniformis.
XX
XX WO200229113-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-US031437.
XX
XX 06-OCT-2000; 2000US-00680598.
XX
XX 27-MAR-2001; 2001US-0279526P.
XX

```

PA (NOVO ) NOVOZYMES BIOTECH INC.  
 XX (NOVO ) NOVOZYMES AS.  
 PI Berka R, Clausen IC;  
 XX WPI; 2002-416684/44.  
 DR  
 XX Monitoring differential expression of several genes in first Bacillus  
 PT cell relative to expression of same genes in one or more second Bacillus  
 PT cells, by using substrate containing Bacillus genomic sequenced tag  
 PT array.  
 XX  
 PS Claim 4; SEQ ID NO 234; 200pp; English.  
 XX  
 CC The invention describes a method of monitoring differential expression of  
 CC genes in a first Bacillus cell relative to expression of the genes in  
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
 CC isolated from Bacillus cells to a substrate containing array of Bacillus  
 CC genomic sequenced tags (GST), examining the array, and determining  
 CC relative gene expression by an observed hybridisation reporter signal of  
 CC a spot in the array. The method is useful for measuring the expression of  
 CC genes in a first Bacillus cell relative to expression of the same genes  
 CC in one or more second Bacillus cells. The method is useful for monitoring  
 CC global expression of several genes from a Bacillus cell, discovering new  
 CC genes, identifying possible functions of unknown open reading frames and  
 CC monitoring gene copy number variation and stability. Monitoring changes  
 CC in expression of genes may be used to provide a representation of the way  
 CC in which Bacillus cells adapt to changes in culture conditions.  
 CC environmental stress or other physiological provocation. Extensive follow  
 CC -up characterisation is unnecessary, when one spot on an array equals one  
 CC gene or one open reading frame, since sequence information is available.  
 CC This sequence represents a genomic sequence tag (GST) used in the method  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1362 BP; 382 A; 334 C; 361 G; 285 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.3; DB 6; Length 1362;  
 Best Local Similarity 89.5%; Pred. No. 7.2e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GGAAGAGGAGGCGCTGTC 20  
 DB 84 GGAAGAGGAGGCGCTGTC 102  
 RESULT 13  
 AAS85172  
 ID AAS85172 standard; cDNA; 1686 BP.  
 XX  
 AC AAS85172;  
 XX  
 XX 13-FEB-2002 (first entry)  
 DT  
 DE DNA encoding novel human diagnostic protein #20976.  
 XX  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-JS08631.  
 XX  
 XX 31-MAR-2000; 2000US-00540217.  
 PR  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX

PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR P-PSDB; ABG20985.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 20976; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1686 BP; 445 A; 397 C; 391 G; 452 T; 0 U; 1 Other;  
 Query Match 79.0%; Score 15.8; DB 5; Length 1686;  
 Best Local Similarity 89.5%; Pred. No. 7.2e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGAAGGAGAGGCGTCGT 19  
 DB 376 AGGAAGGAGAGGCGTCGT 394  
 RESULT 14  
 ADA73217  
 ID ADA73217 standard; DNA; 2000 BP.  
 XX  
 AC ADA73217;  
 XX  
 XX 20-NOV-2003 (first entry)  
 DT  
 DE Rice gene, SEQ ID 6543.  
 XX  
 XX Plant; bacterial infection; fungal infection; viral infection; rice;  
 XX gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 PN WO200300098-A1.  
 XX  
 PD 03-JAN-2003.  
 XX  
 XX 22-JUN-2001; 2001WO-IB001105.  
 PF  
 XX 22-JUN-2001; 2001WO-IB001105.  
 PR  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX





Db 1259 GCAAAGGGGAGGCGTCGTC 1241

RESULT 17  
ADD18763/c  
ID ADD18763 standard; DNA; 3213 BP.  
XX AC ADD18763;  
XX DT 15-JAN-2004 (first entry)  
XX DE Human disease related protein DNA sequence SeqID194.  
XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnery; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing; gene; ds.  
XX OS Homo sapiens.  
XX WO2003018621-A2.  
XX PN 06-MAR-2003.  
XX PD 23-AUG-2002; 2002WO-GB003592.  
XX PF 23-AUG-2001; 2001GB-00020558.  
XX PR 05-OCT-2001; 2001GB-00024037.  
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
XX WPI; 2003-290045/28.  
XX DR P-PSDB; ADD18762.  
XX PT New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX PS Claim 27; SEQ ID NO 194; 424bp; English.  
XX CC This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory,  
CC ophthalmological, antiarteriosclerotic or vulnery activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX SQ Sequence 3213 BP; 768 A; 751 C; 780 G; 914 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 9; Length 3213;  
Best Local Similarity 89.5%; Pred. No. 7.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GCAAAGGAGGCGTCGTC 20  
Db 1259 GCAAAGGGGAGGCGTCGTC 1241

RESULT 18  
AAS75476/c  
ID AAS75476 standard; cDNA; 3229 BP.  
XX AC AAS75476;  
XX DT 13-FEB-2002 (first entry)  
XX DE DNA encoding novel human diagnostic protein #11280.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX WO200175067-A2.  
XX PN 11-OCT-2001.  
XX PD 30-MAR-2001; 2001WO-US008631.  
XX PF 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX DR P-PSDB; ABG11289.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX PS Claim 1; SEQ ID NO 11280; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94584 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 3229 BP; 768 A; 751 C; 784 G; 925 T; 0 U; 1 Other;  
Query Match 79.0%; Score 15.8; DB 5; Length 3229;  
Best Local Similarity 89.5%; Pred. No. 7.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GCAAAGGAGGCGTCGTC 20  
Db 1259 GCAAAGGGGAGGCGTCGTC 1241

RESULT 19  
AAS75478/c

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:48:43 ; Search time 122.706 Seconds  
(without alignments)  
729.318 Million cell updates/sec

Title: US-09-877-819B-35

Perfect score: 20

Sequence: 1 aggaagagagagcgctgc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq2:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq3:\*
- 17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	10	US-09-877-819B-5
C 2	20	100.0	20	10	US-09-877-819B-6
C 3	20	100.0	20	10	US-09-877-819B-33
C 4	20	100.0	20	10	US-09-877-819B-35
C 5	16.8	84.0	48436	9	US-09-927-602-38
C 6	16.4	82.0	680	13	US-10-425-114-15395
C 7	16.4	82.0	37923	13	US-10-087-192-1891
C 8	15.8	79.0	307	10	US-09-910-082A-309
C 9	15.8	79.0	402	14	US-10-079-623-208
C 10	15.8	79.0	1007	13	US-10-425-114-27257
C 11	15.8	79.0	1207	13	US-10-424-599-99564
C 12	15.8	79.0	1362	9	US-09-974-300-234
C 13	15.8	79.0	1588	13	US-10-425-114-1449
C 14	15.8	79.0	7690	15	US-10-311-455-1095
C 15	15.8	79.0	7690	15	US-10-311-455-1095

Sequence 198, Appl  
Sequence 22, Appl  
Sequence 1057, Ap  
Sequence 226, App  
Sequence 3539, Ap  
Sequence 3539, Ap  
Sequence 130645,  
Sequence 130645,  
Sequence 1907, Ap  
Sequence 188, App  
Sequence 1071, Ap  
Sequence 908, App  
Sequence 187, App  
Sequence 3, Appli  
Sequence 174961,  
Sequence 174961,  
Sequence 2, Appli  
Sequence 299, App  
Sequence 66, Appl  
Sequence 191193,  
Sequence 191193,  
Sequence 26197, A  
Sequence 956, App  
Sequence 228954,  
Sequence 228955,  
Sequence 228955,  
Sequence 228955,  
Sequence 14186, A  
Sequence 36404, A  
Sequence 1347, A  
Sequence 32, Appl  
Sequence 159, App  
Sequence 1856, Ap  
Sequence 4097, Ap  
Sequence 20077, A  
Sequence 14081, A  
Sequence 30, Appl  
Sequence 1, Appli  
Sequence 27, Appl  
Sequence 7, Appli  
Sequence 574, App  
Sequence 1326, Ap  
Sequence 1126, Ap  
Sequence 13, Appl  
Sequence 77513, A  
Sequence 3711, Ap  
Sequence 28, Appl  
Sequence 1929, Ap  
Sequence 1935, Ap  
Sequence 40, Appl  
Sequence 205, App  
Sequence 9, Appli  
Sequence 1, Appli  
Sequence 76212, A  
Sequence 76212, A  
Sequence 174763,  
Sequence 174763,  
Sequence 2058, Ap  
Sequence 1, Appli  
Sequence 57819, A  
Sequence 1711, Ap  
Sequence 1, Appli  
Sequence 46066, A  
Sequence 19957, A  
Sequence 3858, Ap  
Sequence 115572,  
Sequence 126517,  
Sequence 35548, A  
Sequence 35548, A  
Sequence 5994, Ap

```

C 88 14.8 74.0 51.2 13 US-10-027-632-194334 Sequence 184334,
C 89 14.8 74.0 51.2 13 US-10-027-632-194335 Sequence 194335,
C 90 14.8 74.0 51.2 16 US-10-027-632-194334 Sequence 194334,
C 91 14.8 74.0 51.2 16 US-10-027-632-194335 Sequence 194335,
C 92 14.8 74.0 52.4 9 US-09-974-300-7215 Sequence 7215, Ap
C 93 14.8 74.0 52.5 13 US-10-424-599-41533 Sequence 41533, A
C 94 14.8 74.0 531 15 US-10-029-386-6226 Sequence 6226, Ap
C 95 14.8 74.0 531 13 US-10-027-632-320214 Sequence 320214,
C 96 14.8 74.0 598 16 US-10-027-632-320214 Sequence 320214,
C 97 14.8 74.0 606 13 US-10-027-632-81728 Sequence 81728, A
C 98 14.8 74.0 606 16 US-10-027-632-81728 Sequence 81728, A
C 99 14.8 74.0 621 9 US-09-974-300-537 Sequence 537, App
C 100 14.8 74.0 650 13 US-10-027-632-260739 Sequence 260739,
C 101 14.8 74.0 650 13 US-10-027-632-260740 Sequence 260740,
C 102 14.8 74.0 650 13 US-10-027-632-260741 Sequence 260741,
C 103 14.8 74.0 650 16 US-10-027-632-260742 Sequence 260742,
C 104 14.8 74.0 650 16 US-10-027-632-260739 Sequence 260739,
C 105 14.8 74.0 650 16 US-10-027-632-260740 Sequence 260740,
C 106 14.8 74.0 650 16 US-10-027-632-260741 Sequence 260741,
C 107 14.8 74.0 650 16 US-10-027-632-260742 Sequence 260742,
C 108 14.8 74.0 788 13 US-10-425-114-25185, A Sequence 25185, A
C 109 14.8 74.0 820 13 US-10-027-632-168542 Sequence 168542,
C 110 14.8 74.0 820 13 US-10-027-632-168543 Sequence 168543,
C 111 14.8 74.0 820 16 US-10-027-632-168542 Sequence 168542,
C 112 14.8 74.0 820 16 US-10-027-632-168543 Sequence 168543,
C 113 14.8 74.0 957 13 US-10-027-632-256060 Sequence 256060,
C 114 14.8 74.0 957 13 US-10-027-632-256061 Sequence 256061,
C 115 14.8 74.0 957 16 US-10-027-632-256062 Sequence 256062,
C 116 14.8 74.0 957 16 US-10-027-632-256061 Sequence 256061,
C 117 14.8 74.0 957 16 US-10-027-632-256062 Sequence 256062,
C 118 14.8 74.0 957 16 US-10-027-632-120606 Sequence 120606,
C 119 14.8 74.0 960 16 US-10-027-632-120606 Sequence 120606,
C 120 14.8 74.0 1167 13 US-10-425-114-24650, A Sequence 24650, A
C 121 14.8 74.0 1539 16 US-10-369-493-42989 Sequence 42989, A
C 122 14.8 74.0 1696 16 US-10-310-154-302 Sequence 302, App
C 123 14.8 74.0 1754 15 US-10-128-714-7358 Sequence 7358, Ap
C 124 14.8 74.0 1884 13 US-10-282-1228-18944 Sequence 18944, A
C 125 14.8 74.0 2256 13 US-10-425-114-34856 Sequence 34856, A
C 126 14.8 74.0 2595 16 US-10-369-493-30492 Sequence 30492, A
C 127 14.8 74.0 2821 16 US-10-104-047-396 Sequence 396, App
C 128 14.8 74.0 3479 10 US-09-758-036-4 Sequence 4, Appli
C 129 14.8 74.0 3480 9 US-09-735-995-1 Sequence 1, Appli
C 130 14.8 74.0 3936 10 US-09-764-891-9807 Sequence 9807, Ap
C 131 14.8 74.0 3950 3 US-09-735-995-3 Sequence 3, Appli
C 132 14.8 74.0 4070 13 US-09-795-651-59 Sequence 59, Appl
C 133 14.8 74.0 4605 13 US-10-194-163-128 Sequence 128, App
C 134 14.8 74.0 5401 10 US-09-764-891-9805 Sequence 9805, Ap
C 135 14.8 74.0 7694 10 US-09-764-891-6232 Sequence 6232, Ap
C 136 14.8 74.0 7694 15 US-10-205-428-550 Sequence 550, App
C 137 14.8 74.0 28932 13 US-10-087-192-757 Sequence 757, App
C 138 14.8 74.0 33239 9 US-09-814-950-3 Sequence 3, Appli
C 139 14.8 74.0 77478 13 US-10-087-192-1378 Sequence 1378, Ap
C 140 14.8 74.0 89328 10 US-09-873-367C-332 Sequence 332, App
C 141 14.8 74.0 126266 13 US-10-087-192-1132 Sequence 1132, Ap
C 142 14.8 74.0 132762 10 US-09-954-556-17 Sequence 17, Appl
C 143 14.8 74.0 194883 13 US-10-087-192-826 Sequence 826, App
C 144 14.8 74.0 302250 9 US-09-952-832-154 Sequence 154, App
C 145 14.8 74.0 659158 9 US-09-771-208-20 Sequence 20, Appl
C 146 14.8 74.0 2940917 13 US-10-027-632-174763 Sequence 174763,
C 147 14.8 74.0 2940917 16 US-10-027-632-174763 Sequence 174763,
C 148 14.8 74.0 412 9 US-09-867-701-732 Sequence 732, App
C 149 14.8 74.0 492 13 US-10-424-599-138374 Sequence 138374,
C 150 14.8 74.0 539 13 US-10-425-114-145 Sequence 145, App
C 151 14.8 74.0 539 13 US-10-027-632-281336 Sequence 281336,
C 152 14.8 74.0 550 16 US-10-027-632-281336 Sequence 281336,
C 153 14.8 74.0 596 9 US-09-864-761-6972 Sequence 6972, Ap
C 154 14.8 74.0 602 13 US-10-027-632-114433 Sequence 114433,
C 155 14.8 74.0 602 16 US-10-027-632-223620 Sequence 223620,
C 156 14.8 74.0 622 13 US-10-027-632-223620 Sequence 223620,
C 157 14.8 74.0 622 16 US-10-027-632-223620 Sequence 223620,
C 158 14.8 74.0 632 13 US-10-027-632-287430 Sequence 287430,
C 159 14.8 74.0 632 16 US-10-027-632-287430 Sequence 287430,
C 160 14.8 74.0 632 16 US-10-027-632-287430 Sequence 287430,

```

## ALIGNMENTS

## RESULT 1

```

US-09-877-819B-5
; Sequence 5, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address tag
US-09-877-819B-5

```

```

Query Match 100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGGAAAGGAGGCGTCGTC 20
    |||||
Db 1 AGGAAAGGAGGCGTCGTC 20
    |||||

```

## RESULT 2

```

US-09-877-819B-6/c
; Sequence 6, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: JS/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

```

```
; OTHER INFORMATION: Capture tag
US-09-877-819B-6

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGTC 20
    |||||
Db 20 AGGAAAGGAGGCGTCGTC 1

RESULT 3
US-09-877-819B-33/c
; Sequence 33, Application US/098778193
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Capture sequence
US-09-877-819B-33

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGTC 20
    |||||
Db 20 AGGAAAGGAGGCGTCGTC 1

RESULT 4
US-09-877-819B-35
; Sequence 35, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address sequence
US-09-877-819B-35

Query Match      100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGTC 20
    |||||
Db 1 AGGAAAGGAGGCGTCGTC 20

RESULT 5
US-09-877-819B-35
```

```
US-09-927-602-38/c
; Sequence 38, Application US/099276C2
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-IJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 48436
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-927-602-38

Query Match      84.0%; Score 16.8; DB 9; Length 48436;
Best Local Similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGTC 20
    |||||
Db 2137 AGGAAAGGAGGCGTCATC 2118

RESULT 6
US-10-425-114-15395
; Sequence 15395, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15395
; LENGTH: 880
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-048-G11_FLI
US-10-425-114-15395

Query Match      82.0%; Score 16.4; DB 13; Length 880;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGTC 18
    |||||
Db 47 AGGAAAGGAGGCGGCG 64

RESULT 7
US-10-087-192-1891/c
; Sequence 1891, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
```



```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99564
; LENGTH: 1207
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60921C.1
US-10-424-599-99564

Query Match      79.0%; Score 15.8; DB 13; Length 1207;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGT 19
    |||||
DB 45 AGGAAAGGAGGCGTCGT 27
    |||||

RESULT 12
US-09-974-300-234
; Sequence 234, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085-500-JS
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 234
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-234

Query Match      79.0%; Score 15.8; DB 9; Length 1362;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGAGGAGGCGTCGT 20
    |||||
DB 84 GGAAGAGGAGGCGTCGT 102
    |||||

RESULT 13
US-10-425-114-1449
; Sequence 1449, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1449
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Zea mays

; FEATURE:
; OTHER INFORMATION: Clone ID: 700153483_FLI
US-10-425-114-1449

Query Match      79.0%; Score 15.8; DB 13; Length 1588;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGAGGAGGCGTCGT 20
    |||||
DB 135 GGAAGAGGAGGCGTCGT 153
    |||||

RESULT 14
US-10-311-455-1095
; Sequence 1095, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: PIRPENEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1095
; LENGTH: 7690
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1095

Query Match      79.0%; Score 15.8; DB 15; Length 7690;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGAAAGGAGGCGTCGT 19
    |||||
DB 7531 AGGAAAGGAGGCGTCGT 7549
    |||||

RESULT 15
US-10-311-455-198
; Sequence 198, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIRPENEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 198
; LENGTH: 14924
; TYPE: DNA
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1272.82 Seconds  
(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819B-35

Perfect score: 20

Sequence: 1 aggaagaggagggctgcgtc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

E8T:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	733	28	B2967818 PJGZH24TB
C 2	17.4	87.0	379	9	AB076957 AB076957
C 3	17.4	87.0	503	14	CD111750 ME1-0021T
C 4	17.4	87.0	695	28	BH705590 BOMMH07TR

C	5	17.4	87.0	773	28	BH718652
C	6	16.8	84.0	358	10	BB813656
C	7	16.8	84.0	475	10	BF622509
C	8	16.8	84.0	527	29	LBFA020504
C	9	16.8	84.0	530	28	BZ894876
C	10	16.8	84.0	551	29	CC483605
C	11	16.8	84.0	554	13	BU214159
C	12	16.8	84.0	570	13	EX276312
C	13	16.8	84.0	572	13	BU107606
C	14	16.8	84.0	572	28	BZ349842
C	15	16.8	84.0	577	13	EX276311
C	16	16.8	84.0	591	13	BU108783
C	17	16.8	84.0	619	10	BF224860
C	18	16.8	84.0	651	13	BU289346
C	19	16.8	84.0	665	29	CNS01R25
C	20	16.8	84.0	679	14	CD430028
C	21	16.8	84.0	681	12	BM489672
C	22	16.8	84.0	697	29	LBFA037H08
C	23	16.8	84.0	744	29	CG331413
C	24	16.8	84.0	748	13	BU143208
C	25	16.8	84.0	748	13	BU144044
C	26	16.8	84.0	755	29	CG331399
C	27	16.8	84.0	759	12	BIS21933
C	28	16.8	84.0	763	29	AG075431
C	29	16.8	84.0	765	12	B1831377
C	30	16.8	84.0	771	13	BU144412
C	31	16.8	84.0	771	28	CC393953
C	32	16.8	84.0	780	13	BU246064
C	33	16.8	84.0	781	13	BU109258
C	34	16.8	84.0	783	13	BU144375
C	35	16.8	84.0	790	13	BU145093
C	36	16.8	84.0	813	14	CA764236
C	37	16.8	84.0	815	13	BU144698
C	38	16.8	84.0	829	13	BU144278
C	39	16.8	84.0	841	29	CG289248
C	40	16.8	84.0	872	13	BU143249
C	41	16.8	84.0	885	14	CD241445
C	42	16.8	84.0	929	13	BU106836
C	43	16.8	84.0	960	13	BU144429
C	44	16.8	84.0	962	12	BU269135
C	45	16.8	84.0	983	12	EG343181
C	46	16.8	84.0	996	13	BU144543
C	47	16.8	84.0	1041	13	BU143936
C	48	16.8	84.0	1048	13	BU108727
C	49	16.8	84.0	1050	29	CNS0100D
C	50	16.8	84.0	1057	13	BU506565
C	51	16.8	84.0	1097	29	CNS02AD3
C	52	16.8	84.0	1101	29	CNS05OTX
C	53	16.8	84.0	1594	11	BC032063
C	54	16.4	82.0	285	10	BB184379
C	55	16.4	82.0	580	28	BH569276
C	56	16.4	82.0	641	14	CD296086
C	57	16.4	82.0	705	29	LBFA082C03
C	58	16.4	82.0	784	14	CF748983
C	59	16.4	82.0	785	29	CC605307
C	60	16.4	82.0	786	28	BZ216439
C	61	16.4	82.0	801	28	BH585750
C	62	16.4	82.0	954	12	BG860059
C	63	16.4	82.0	1201	13	EX464371
C	64	16	80.0	209	12	B1336789
C	65	16	80.0	288	10	BB071518
C	66	16	80.0	357	12	BM744826
C	67	16	80.0	371	12	BM737808
C	68	16	80.0	414	14	CF938309
C	69	16	80.0	484	14	CD667452
C	70	16	80.0	561	14	CF943655
C	71	16	80.0	574	28	AZ004718
C	72	16	80.0	778	10	BF346996
C	73	16	80.0	1186	10	BF688771
C	74	15.8	79.0	127	13	BQ535248
C	75	15.8	79.0	183	29	CE439441
C	76	15.8	79.0	186	10	BE600743
C	77	15.8	79.0	217	10	BF651869

BH718652	BONMM90TR
BB813656	BB813656
BF622509	HVSMES000
EX539720	Leishmani
BZ894876	H94_0.66
CC483605	CH240_312
BU214159	603109140
EX276312	BX276312
BU107606	603111219
BZ349842	h-46d12_9
EX276311	BX276311
BU108783	603109972
BF224860	uz11g12_x
BU289346	604165581
AL163994	Tetraodon
CD430028	ETH1_13_F
BM489672	pgm2n.pk0
EX59026	Leishmani
CG331413	OGXDG51TV
BU143208	603162366
BU144044	603228840
CG331399	OGXDG51TH
BH521933	603081256
AG075431	pan trogi
B1831377	603074366
BU144412	603229615
CC393953	PURKG45TB
BU346064	604172642
BU109258	603112171
BU144375	603229576
BU145093	603229423
CA764236	AF53-Rp1
BU144698	603229629
BU144278	603229808
CG289248	OGXDI83TV
BU143249	603229230
CD241445	AGENCOURT
BU106836	603110686
BU144429	603230184
BU269135	603598625
BG343181	HVSMES000
BU144543	603228913
BU143936	603230222
BU108727	603112393
AL098359	Drosophil
BU506565	AGENCOURT
AL184400	Tetraodon
AL347118	Tetraodon
BC032063	Homo sapi
BB184379	BB184379
BH569276	BOQV37TF
CD296086	StrPu691
BX38515	Leishmani
CF748983	UI-M-HU0-
CC605307	OGVFP16TV
BZ216439	CH230-323
BH585750	BOGFL84TF
BG860059	1024067B1
EX464371	BX464371
B1336789	AR081D12B
BB071518	BB071518
BM744826	K-EST0018
BM737808	K-EST0001
CF938309	NC-ESTcab3
CD667452	NC-EST3b92
CF943655	NC-ESTqac1
AZ004718	RPCI-23-3
BF346996	602021969
BF688771	602184936
BQ535248	L2AF3_9_F
CE439441	tigr-985-
BE600743	P1189_F1
BF651869	274762_MA



```

c 78 15.8 79.0 220 9 AV215370
c 79 15.8 79.0 235 14 D47401
c 80 15.8 79.0 239 10 BE598721 P11_81_C0
c 81 15.8 79.0 260 10 AW327123 20682 MAR
c 82 15.8 79.0 271 10 AW432033 73701 MAR
c 83 15.8 79.0 278 14 CD191397 G608.113D
c 84 15.8 79.0 291 12 BI040676 CX4-NT028
c 85 15.8 79.0 304 29 CG171414 1119221C1
c 86 15.8 79.0 322 13 EQ167794 WHE0004.E
c 87 15.8 79.0 333 14 CD917511 G608.105K
c 88 15.8 79.0 335 14 CB670779 OSJNFe04A
c 89 15.8 79.0 348 14 CF638075 ZmrW00.0
c 90 15.8 79.0 353 26 BZ332999 bx32b03.G
c 91 15.8 79.0 358 10 BE232041 136954 MA
c 92 15.8 79.0 364 13 BY173691 BX173691
c 93 15.8 79.0 366 12 BM660035 EST042.Z
c 94 15.8 79.0 393 13 BY167002 tigr-gss-
c 95 15.8 79.0 395 23 BE694400 tigr-gss-
c 96 15.8 79.0 400 14 CD736027 4017405.1
c 97 15.8 79.0 401 13 CA117950 SCBGR102
c 98 15.8 79.0 401 13 CA123893 SCNCR112
c 99 15.8 79.0 407 13 CA828524 1114029C0
c 100 15.8 79.0 410 13 BY627189 BY627189
c 101 15.8 79.0 414 13 BY538356 BX538356
c 102 15.8 79.0 428 28 AQ922952 R3CI-23-2
c 103 15.8 79.0 432 13 CA027001 H257117r
c 104 15.8 79.0 437 29 CC556911 ZMMBSC039
c 105 15.8 79.0 438 14 CF642384 D51_A08.F
c 106 15.8 79.0 444 13 BY535664 BY535664
c 107 15.8 79.0 446 9 AU223334 AU223334
c 108 15.8 79.0 448 28 AQ594450 HS.2105.B
c 109 15.8 79.0 451 13 BY530097 BY530097
c 110 15.8 79.0 451 28 AQ608036 HS.2128.B
c 111 15.8 79.0 452 28 AZ048267 LMAJFV1.1
c 112 15.8 79.0 452 29 CG712047 1119023H0
c 113 15.8 79.0 467 13 BU087634 Na.L3.22F
c 114 15.8 79.0 468 14 CD220050 CCG1.60.E
c 115 15.8 79.0 468 29 CE050725 tigr-gss-
c 116 15.8 79.0 478 14 CZ135231 UT-HP-CB0
c 117 15.8 79.0 485 9 AW099926 sdlsae02.Y
c 118 15.8 79.0 486 10 AW563353 LGI.235.G
c 119 15.8 79.0 486 10 BP252125 EST419387
c 120 15.8 79.0 491 12 BG463617 EM1.50.G1
c 121 15.8 79.0 495 9 AA432497 AA432497
c 122 15.8 79.0 495 28 BH385788 AG-ND-167
c 123 15.8 79.0 501 12 EG649746 RI122.83
c 124 15.8 79.0 503 28 AZ004120 RPCI-23-3
c 125 15.8 79.0 504 13 EX564726 BX564726
c 126 15.8 79.0 505 12 BM499068 949015G09
c 127 15.8 79.0 505 14 CF512940 CF512940
c 128 15.8 79.0 509 9 AV429732 AV429732
c 129 15.8 79.0 510 10 BE752684 255143.MA
c 130 15.8 79.0 510 13 BU711808 BU711808
c 131 15.8 79.0 512 13 BX568933 BX568933
c 132 15.8 79.0 515 14 CD058826 3529.1.10
c 133 15.8 79.0 516 9 AA108379 AA108379
c 134 15.8 79.0 516 29 TA201A10Q
c 135 15.8 79.0 517 13 BY471875 BY471875
c 136 15.8 79.0 517 14 R41658 R41658
c 137 15.8 79.0 518 13 BQ485312 BQ485312
c 138 15.8 79.0 519 10 BE600821 P11.90.E1
c 139 15.8 79.0 519 28 CC018569 3591.1.10
c 140 15.8 79.0 519 28 AQ660736 Shear3.D
c 141 15.8 79.0 522 13 BU997493 BU997493
c 142 15.8 79.0 523 12 EI098288 EI130.E0
c 143 15.8 79.0 526 14 CD290518 StrPu538.
c 144 15.8 79.0 526 28 AQ118411 HS.3007.A
c 145 15.8 79.0 528 10 BE359311 LGI.40.E0
c 146 15.8 79.0 528 29 CE492180 tigr-gss-
c 147 15.8 79.0 529 14 CD318639 StrPu538.
c 148 15.8 79.0 535 10 BF421362 FMI.7.G01
c 149 15.8 79.0 537 13 BU722286 SJMAFF04
c 150 15.8 79.0 538 10 AW672099 LGI.357.H

```

```

151 15.8 79.0 539 12 BI074696
152 15.8 79.0 543 12 BM326188
153 15.8 79.0 547 12 BI013499
154 15.8 79.0 554 10 BE123288
155 15.8 79.0 554 14 CA826537 1114001G0
156 15.8 79.0 559 14 CA263490 SCRLB203
157 15.8 79.0 563 14 CA300899 SCRLJ105
158 15.8 79.0 563 13 BU986337 HFL1H10r
159 15.8 79.0 565 13 CA278891 SCBFT309
160 15.8 79.0 571 14 CF432984 NIT1_20_C
161 15.8 79.0 571 14 CB927140 ABAT_13_G
162 15.8 79.0 574 10 BF787850 F02113629
163 15.8 79.0 574 10 AW562579 660066F04
164 15.8 79.0 575 28 BZ625814 lb41a06.G
165 15.8 79.0 581 24 CA575352 K062/G12
166 15.8 79.0 581 29 CE619857 tigr-gss-
167 15.8 79.0 585 14 CA299429 SCRLFL805
168 15.8 79.0 590 14 CD827405 BN25.067F
169 15.8 79.0 590 14 CF873121 trico04Xa
170 15.8 79.0 590 28 AQ660152 Shear3.D
171 15.8 79.0 595 14 CF295331 30DGS-05
172 15.8 79.0 596 13 CA070867 SCUTAD103
173 15.8 79.0 597 14 CA193406 SCCPL100
174 15.8 79.0 598 10 AW745490 WS1.35.D0
175 15.8 79.0 601 28 AZ082435 RPCI-23-4
176 15.8 79.0 602 12 BM439965 pgr.in.pko
177 15.8 79.0 604 14 CA242680 SCSPFL309
178 15.8 79.0 605 14 CA786488 SCSPFL106
179 15.8 79.0 607 12 EM076459 TREST-A33
180 15.8 79.0 608 14 CD463990 ETH1_47_A

```

## ALIGNMENTS

```

RESULT 1
LOCUS BZ967818/c 733 bp DNA linear GSS 25-MAR-2003
DEFINITION PUGH2147B.ZM.0.6.1.0.KB.Zea.mays.genomic.clone.ZMMBTA388D04,
genomic survey sequence.
ACCESSION BZ967818
VERSION BZ967818.1 GI:29185956
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 733)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUGH2147D
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
location/Qualifiers
1..733
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA388D04"
/clone_lib="ZM.0.6.1.0.KB"
/notes=vector: PCR4-TOPO; site:1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

```

ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 733;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 AGGAAGGAGGAGGCGTCGTC 20  
|||||  
Db 725 AGGAAGGAGGAGGCGTCGTC 706

RESULT 2

AB076957/c  
LOCUS  
DEFINITION AB076957 Human vestibular cDNA library Homo sapiens cDNA clone  
402V5-12-54, mRNA sequence.  
ACCESSION AB076957  
VERSION AB076957.1 GI:21678535  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 379)  
Abe, S., Koyama, K., Usami, S. and Nakamura, Y.  
TITLE Construction of a vestibular-specific cDNA library  
JOURNAL Unpublished (2002)  
COMMENT Contact: Satoko Abe  
Institute of Medical Science  
The University of Tokyo, Human Genome Center  
4-6-1, Minato-ku, Tokyo 106-8639, Japan  
Tel: 81-3-5449-5375  
Fax: 81-3-5449-5406  
Email: satoko@ims.u-tokyo.ac.jp.  
Location/Qualifiers  
1. 379  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="402V5-12-54"  
/tissue\_type="vestibule"  
/clone\_lib="Human vestibular cDNA library"

FEATURES

source  
1. 379  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="402V5-12-54"  
/tissue\_type="vestibule"  
/clone\_lib="Human vestibular cDNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 379;  
Best Local Similarity 94.7%; Pred. No. 3.2e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 AGGAAGGAGGAGGCGTCGTC 19  
|||||  
Db 37 AGGAAGGAGGAGGCGTCGTC 19

RESULT 3

CD111750/c  
LOCUS  
DEFINITION ME1-0021T-D051-E02-U.B ME1-0021 Schistosoma mansoni cDNA clone  
ME1-0021T-D051-E02.B, mRNA sequence.  
ACCESSION CD111750  
VERSION CD111750.1 GI:34649868  
KEYWORDS EST.  
SOURCE Schistosoma mansoni  
ORGANISM  
Schistosoma mansoni  
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.  
REFERENCE 1 (bases 1 to 503)  
AUTHORS  
Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,  
Ojori, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,  
Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,  
Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,  
Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,  
Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,  
Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,

TITLE

JOURNAL  
MEDLINE  
COMMENT

Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,  
Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.  
Transcriptome analysis of the acelomate human parasite Schistosoma  
mansoni  
Nat. Genet. 35 (2), 148-157 (2003)  
22879926  
Contact: Dr. Sergio Verjovski-Almeida  
Departamento de Bioquímica  
Instituto de Química - Universidade de São Paulo  
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
Brasil  
Tel: +55-11-3091-2173  
Fax: +55-11-3091-2186  
Email: verjo@iq.usp.br  
This sequence was derived from the FAPESP Schistosoma mansoni EST  
Genome Project. All sequences in the project were assembled and  
annotated. This entry and all the assembled sequences can be seen  
in the following URL <http://bioinfo.iq.usp.br/schisto/>  
Plate: ME1-0021T-D051 row: 2 column: E.

FEATURES

source  
1. 503  
/organism="Schistosoma mansoni"  
/mol\_type="mRNA"  
/db\_xref="taxon:6183"  
/clone="ME1-0021T-D051-E02.B"  
/sex="mixed pool"  
/dev stage="egg"  
/lab\_host="Mus musculus"  
/clone\_lib="ME1-0021"

ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 503;  
Best Local Similarity 94.7%; Pred. No. 3.5e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 GGAAAGGAGGAGGCGTCGTC 20  
|||||  
Db 440 GGCAAGGAGGAGGCGTCGTC 422

RESULT 4

BH705590/c  
LOCUS  
DEFINITION BOMMH07T BO.2.3 KB Brassica oleracea genomic clone BOMMH07,  
genomic survey sequence.  
ACCESSION BH705590  
VERSION BH705590.1 GI:18788415  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 695)  
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other GSSs: BOMMH07TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1. 695  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="Tol000DH3"  
/db\_xref="taxon:3712"

FEATURES

source

```

/clone="BOMMH07"
/clone_lib="BO_2_3_KB"
/notes=vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      87.0%; Score 17.4; DB 28; Length 695;
Best Local Similarity 94.7%; Pred. No. 3.7e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAAAGGAGGAGCGTCGTC 20
    |||||
DB 23 GGAAAGGAGGAGCGTCGTC 5

RESULT 5
BH718652/c
LOCUS BH718652 773 bp DNA linear GSS 20-FEB-2002
DEFINITION BOMM90TR BO_2_3_KB Brassica oleracea genomic clone BOMM90,
genomic survey sequence.
ACCESSION BH718652
VERSION BH718652.1 GI:18817384
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 773)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOMM90TF
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: IR
Class: sheared ends.
Location/Qualifiers
1..773
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
/db_xref="taxon:3712"
/clone="BOMM90"
/clone_lib="BO_2_3_KB"
/notes=vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

FEATURES
source
1..773
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
/db_xref="taxon:3712"
/clone="BOMM90"
/clone_lib="BO_2_3_KB"
/notes=vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      87.0%; Score 17.4; DB 28; Length 773;
Best Local Similarity 94.7%; Pred. No. 3.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAAAGGAGGAGCGTCGTC 20
    |||||
DB 482 GGAAAGGAGGAGCGTCGTC 464

RESULT 6
BB813656
LOCUS BB813656 358 bp mRNA linear EST 19-NOV-2001
DEFINITION BB813656 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus
musculus cDNA clone G730020C03 3', mRNA sequence.
ACCESSION BB813656
VERSION BB813656.1 GI:16986285
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 358)
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hiracka, T., Hirozane, T., Imotani, K.,
Ishii, Y., Ito, M., Kawai, O., Kojima, Y., Konno, H., Kouda, M.,
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: http://genome-gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome-gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1..358
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G730020C03"
/tissue_type="lung"
/cell_line="RCB-0558 LLC"
/clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
cDNA"

ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 358;
Best Local Similarity 90.0%; Pred. No. 5.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGAAGGAGGAGCGTCGTC 20
    |||||
DB 63 AGGAAGGAGGAGCGTCGTC 82

RESULT 7
BF622509/c
LOCUS BF622509 475 bp mRNA linear EST 17-OCT-2001
DEFINITION BF622509 HVSMEa0004K15f Hordeum vulgare seedling shoot EST library
HVCNDA0001 (Cold stress) Hordeum vulgare subsp. vulgare cDNA clone
HVSMEa0004K15f, mRNA sequence.
ACCESSION BF622509
VERSION BF622509.2 GI:13081086

```

## KEYWORDS

SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.

1 (bases 1 to 475)

Wing,R., Close,T.J., Kleinohs,A., Wise,R., Begum,D., Frisch,D.,  
Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R.,  
Choi,D.W., Fenton,R.D. and Main,D.

Development of a genetically and physically anchored EST resource  
for barley genomics: Morex cold-stressed seedling shoot cDNA  
library

Unpublished (2001)

On Dec 18, 2000 this sequence version replaced gi:11886243.

Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 271

Seq primer: AATTAACCTCACTAAAGGG

High quality sequence stop: 467.

## FEATURES

Location/Qualifiers

1..475  
/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Morex"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="HVSMEa0004K15f"

/tissue\_type="Seedling shoot"

/lab\_host="TJCI121"

/clone\_lib="Hordeum vulgare seedling shoot EST library"

HVCDA0001 (Cold stress)"

/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
Seeds were surface sterilized then germinated under axenic  
conditions in the dark at room temperature on filter paper  
with water, nystatin and cefotaxime in covered  
crystallization dishes. Five-day old seedlings were  
incubated at 50C for 2 days. Shoots were then harvested,  
total RNA was prepared, poly(A) RNA was purified, one  
primary unamplified cDNA library was made, and 60000 pfu  
were in vivo excised to give pBluescript SK(-) cDNA  
phagemids. These steps were performed in the TJ Close  
laboratory at the University of California, Riverside  
(Choi, Close, Fenton). Phagemids were plated and picked at  
the Clemson University Genomics Institute (CUGI) (Begum,  
Palmer, Frisch, Atkins and Wing). Plasmid DNA  
preparations, DNA sequencing and sequence analysis were  
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,  
Oates, Rambo, Main). The sequence has been trimmed to  
remove vector sequence and contains a minimum of 100 bases  
of phred value 20 or above. For more details on library  
preparation and sequence analysis see  
<http://www.genome.clemson.edu/projects/barley>. To order  
this clone see <http://www.genome.clemson.edu/orders> Also  
see Close TJ, Wing R, Kleinohs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(<http://wheat.pw.usda.gov/g9ppages/bgn/31/cover.html>)"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 475;  
Best Local Similarity 90.0%; Pred. No. 6.1e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAACGAGGCGTCGTC 20

||||| ||||||| |||||

Db 62 AGGAACGAGGCGTCGTC 43

## RESULT 8

LBAP090E04

LOCUS

DEFINITION

Leishmania braziliensis GSS, clone LBAP090E04, genomic survey

sequence.

ACCESSION

BX539720

VERSION

1 GI:32138920

KEYWORDS

GSS; Genomic survey sequence.

SOURCE

Leishmania braziliensis

ORGANISM

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Leishmania; Leishmania braziliensis species complex.

REFERENCE

2 (bases 1 to 527)

AUTHORS

Cruz,A.K.

TITLE

Direct Submission

JOURNAL

Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo,

Department of Molecular and Cell Biology, FMRP, Avenida

Bandeirantes 3906, Ribeirao Preto, SP, 14049-900, BRAZIL

Clone requests: akcruz@fmrp.usp.br.

COMMENT

Location/Qualifiers

1..527

/organism="Leishmania braziliensis"

/mol\_type="genomic DNA"

/strain="MHOM/BR/75/M2904"

/db\_xref="taxon:5660"

/clone="LBAP090E04"

ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 527;

Best Local Similarity 90.0%; Pred. No. 6.3e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAGGAGGCGTCGTC 20

||||| ||||||| |||||

Db 358 AGGAGGAGGCGTCGTCG 377

## RESULT 9

BZ894876

LOCUS

DEFINITION

Hg4.0166 Hg pUC18 Library Halobaculum gomorrense genomic 5',

genomic survey sequence.

ACCESSION

BZ894876

VERSION

BZ894876.1 GI:33345352

KEYWORDS

GSS.

SOURCE

Halobaculum gomorrense

ORGANISM

Halobaculum gomorrense

REFERENCE

1 (bases 1 to 530)

AUTHORS

Goo,Y., Roach,J., Glusman,G., Baliga,N.S., Deutsch,K., Pan,M.,

Dassarma,S., Ng,W.V. and Hood,L.

TITLE

Low-pass Sequencing for Microbial Comparative Genomics

JOURNAL

Unpublished (2003)

COMMENT

Contact: Goo Y

Institute for Systems Biology

1441 North 34th Street, Seattle, WA 98103, USA

Tel: 206 732 1412

Fax: 206 732 1299

Email: ygoo@systemsbiology.org

Seq primer: M13 Forward

Class: shotgun.

FEATURES

Location/Qualifiers

1..530

/organism="Halobaculum gomorrense"

/mol\_type="genomic DNA"

/strain="ATCC 700876"

/db\_xref="taxon:43928"

/clone\_lib="Hg pUC18 Library"  
 /note="Vector: pUC18; Site 1: SmaI; A shotgun library was  
 constructed from Halobaculum gomorrense genomic DNA using  
 pUC18/SmaI/3AP plasmid"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 28; Length 530;  
 Best Local Similarity 90.0%; Pred. No. 6.4e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGAAGGAGAGCGTCGTC 20  
 |||||  
 Db 228 AGGAAGGAGAGCGTCGTC 247  
 |||||

RESULT 10  
 CC483605/c  
 LOCUS CC483605  
 DEFINITION CC483605 551 bp DNA linear GSS 16-JUN-2003  
 CH240\_312D19.77 CHORI-240 Bos taurus genomic clone CH240\_312D19,  
 genomic survey sequence.  
 ACCESSION CC483605  
 VERSION CC483605.1 GI:31765110  
 KEYWORDS GSS  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 551)  
 Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,  
 Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,  
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,  
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,  
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,  
 Dalrymple, B. P. and Tellam, R.  
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398  
 Unpublished (2003)  
 Other GSSs: CH240\_312D19.77BAC13P2  
 Contact: Rob Holt  
 Sequencing  
 The British Columbia Cancer Agency Genome Science Centre  
 600 W. 10th Ave. Vancouver, British Columbia, Canada V5Z 4E6  
 Tel: 604-877-6085  
 Fax: 604-877-6276  
 Email: rholt@bccsc.ca  
 Clones are derived from the bovine BAC library CHORI-240  
 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library  
 availability, please contact Pieter de Jong ([pdejong@mail.choi.org](mailto:pdejong@mail.choi.org)).  
 Clones may be purchased from BACPAC Resources  
 (<http://www.chori.org/bacpac/ordering/information.htm>). This work  
 was undertaken as part of the International Bovine BAC Mapping  
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the  
 British Columbia Genome Sciences Centre, Canada.  
 Seq primer: T7  
 Plate: 312 row: D column: 19  
 Class: BAC ends.

## FEATURES

Location/Qualifiers  
 1..551  
 /organism="Bos taurus"  
 /mol\_type="genomic DNA"  
 /st-ain="breed: Hereford"  
 /db\_xref="taxon:9913"  
 /clone="CH240\_312D19"  
 /sex="Male"  
 /cell\_type="Blood"  
 /clone\_lib="CHORI-240"  
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;  
 Hereford bull 141 Domino 99375; CHORI-240 Bovine BAC  
 library (Male) produced by Pieter de Jong"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 551;  
 Best Local Similarity 90.0%; Pred. No. 6.4e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGAAGGAGAGCGTCGTC 20  
 |||||  
 Db 548 AGGAAGGAGAGCGTCGTC 529  
 |||||

RESULT 11  
 BU214159/c  
 LOCUS BU214159  
 DEFINITION BU214159 554 bp mRNA linear EST 25-NOV-2002  
 603109140F1 CSROCHN04 Gallus gallus cDNA clone CHESY5309 5', mRNA  
 sequence.  
 ACCESSION BU214159  
 VERSION BU214159.1 GI:25389969  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 554)  
 Boardman, P. E., Sanz-Esquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,  
 Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

Location/Qualifiers  
 1..554  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHESY5309"  
 /tissue\_type="whole embryo"  
 /dev\_stage="20-21"  
 /lab\_host="DH10B"  
 /clone\_lib="CSROCHN04"  
 /note="Organ: whole embryo; Vector: pBluescript II KS(+);  
 Site 1: EcoRI; Site 2: NotI; This normalized library was  
 constructed from 1 million independent clones. cDNA  
 synthesis was initiated using an oligo(dT) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunted, ligated to NotI adapters, digested with  
 EcoRI, size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 554;  
 Best Local Similarity 90.0%; Pred. No. 6.4e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGAAGGAGAGCGTCGTC 20  
 |||||  
 Db 164 AGGAAGGAGAGCGTCGTC 145  
 |||||

RESULT 12  
 BX276312/c



Plate: hz46 row: d column: 12  
 Seq primer: -21M13UnivRev  
 Class: Shotgun  
 High quality sequence stop: 572.  
 Location/Qualifiers  
 1..572

## FEATURES

source

```

/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/clone="hr46d12"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"
/notes="Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a."
```

## ORIGIN

Query Match 84.0%; Score 16.8; DB 28; Length 572;  
 Best Local Similarity 90.0%; Pred. No. 6.4e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAGAGGAGGCGTCGTC 20  
 |||||  
 Db 378 AGGAAGAGGAGGCGTCGCC 359  
 |||||

## RESULT 15

EX276311  
 LOCUS  
 DEFINITION BX276311 AGENAE Gallus gallus multi-tissues normalized library  
 (gcag) Gallus gallus cDNA clone gcag0008c.e.09 3prim, mRNA

ACCESSION BX276311  
 VERSION BX276311.1 GI:28598802  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE  
 AUTHORS Herault, F., Le Meuth-Metzinger, V., Desert, C., Retout, E., Piumi, F., Klopp, C. and Douaire, M.  
 TITLE Construction and primary characterization of chicken normalized multi-tissue cDNA libraries

## JOURNAL

Unpublished (2003)

COMMENT

Contact: Douaire M

INRA, UMR INRA-ENSAR Genetique Animale

65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE

Tel: +33 (0) 2.23.48.54.63

Fax: +33 (0) 2.23.48.54.70

Email: Madeleine.Douaire@roazhon.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us at signasupport@jouy.inra.fr to obtain the chromatogram of this sequence

Plate: 0008 row: e column: 9

Seq primer: M13P.

Location/Qualifiers

1..577

/organism="Gallus gallus"

/mol\_type="mRNA"

/db\_xref="taxon:9031"

/clone="gcag0008c.e.09"

/tissue\_type="multi-tissues"

/dev\_stage="from embryos to adults"

/lab\_host="DH10B"

/clone\_lib="AGENAE Gallus gallus multi-tissues normalized library (gcag)"

## FEATURES

source

```

/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcag0008c.e.09"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized library (gcag)"
```

## ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 591;  
 Best Local Similarity 90.0%; Pred. No. 6.5e+03;

/note="Vector: pT73D-pac; tissues: brain, embryos, kidney, multi-tissues, muscle, pancreas, skin, testis, liver, adipose tissue, granulosa, utero-vaginal gland, oviduct, small follicle, ovary, hypothalamus, pituitary gland, ileon, jejunum, caecum, duodenum, spleen, fabricius gland, bone marrow, thymus, hematopoietic progenitor cells. Clone distribution: AGENAE Resource centre, Francois PIUMI, Francois.PIUMI@roazhon.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Couy-en-Josas cedex, FRANCE"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 577;  
 Best Local Similarity 90.0%; Pred. No. 6.4e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGAAGAGGAGGCGTCGTC 20  
 |||||  
 Db 404 AGGAAGAGGAGGAGGTTGTC 423  
 |||||

## RESULT 16

BUI08783/c  
 LOCUS  
 DEFINITION BUI08783 591 bp mRNA linear EST 25-NOV-2002  
 60109972F1 CSEBCHL12 Gallus gallus cDNA clone ChEST55017 5', mRNA sequence.

ACCESSION BUI08783  
 VERSION BUI08783.1 GI:25311874  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE  
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Rickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 22335534  
 PUBMED 12445392

## COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 10D, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..591

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="White Leghorn, Hisex"

/db\_xref="taxon:9031"

/clone="ChEST55017"

/dev\_stage="36"

/lab\_host="DH10B"

/clone\_lib="CSEBCHL12"

/note="Organ: heart; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the t-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BsgI and BamHI sites [5'ggccgctgcagcccgctccgagaaaaag] [5'aattcttttttcgctccgggtgcagc]."

## Query Match

84.0%; Score 16.8; DB 13; Length 591;

## Best Local Similarity

90.0%; Pred. No. 6.5e+03;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 604.235 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819b-36

Perfect score: 20

Sequence: 1 ttaagagcgctccaaagcc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516955 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbi:\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_man.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
C 1	18.4	92.0	158412	2	AP001134	AP001134 Homo sapi
C 2	18	90.0	2144	9	EC034803	EC034803 Homo sapi
C 3	18	90.0	132953	9	AC024221	AC024221 Homo sapi
C 4	18	90.0	161087	2	AC018850	AC018850 Homo sapi
C 5	18	90.0	170067	2	AC124946	AC124946 Homo sapi
C 6	18	90.0	188526	2	AC012342	AC012342 Homo sapi
C 7	18	90.0	190943	2	AC130424	AC130424 Homo sapi
C 8	18	90.0	328187	2	AC117393	AC117393 Homo sapi
C 9	17.4	87.0	216198	2	AC109271	AC109271 Mus muscu
C 10	16.8	84.0	30424	9	AC021658	AC021658 Homo sapi
C 11	16.8	84.0	42775	1	ATVIRB	X06826 Agrobacteri
C 12	16.8	84.0	65825	9	AL161904	AL161904 Human DNA
C 13	16.8	84.0	72440	2	AC113794_5	Continuation (6 of
C 14	16.8	84.0	92800	9	AC004486	AC004486 Homo sapi
C 15	16.8	84.0	110000	2	BX546456_0	BX546456 Homo sapi
C 16	16.8	84.0	119730	9	HSJ989D7	AL121911 Human DNA
C 17	16.8	84.0	121866	2	HSJ322E17	AL109937 Homo sapi
C 18	16.8	84.0	123066	2	AC140029	AC140029 Medicago
C 19	16.8	84.0	136384	2	AL365221	AL365221 Homo sapi
C 20	16.8	84.0	153920	2	AC011723	AC011723 Homo sapi
C 21	16.8	84.0	160659	2	AC036235	AC036235 Homo sapi
C 22	16.8	84.0	163772	9	AC093165	AC093165 Homo sapi
C 23	16.8	84.0	163976	2	AC102004	AC102004 Mus muscu
C 24	16.8	84.0	165935	9	BX537254	BX537254 Human DNA
C 25	16.8	84.0	166098	9	AL356004	AL356004 Human DNA
C 26	16.8	84.0	169788	2	AC113871	AC113871 Rattus no
C 27	16.8	84.0	175588	9	AL139326	AL139326 Human DNA
C 28	16.8	84.0	176182	2	AC137453	AC137453 Rattus no
C 29	16.8	84.0	178505	2	AL611965	AL611965 Homo sapi
C 30	16.8	84.0	183379	2	AC121008	AC121008 Rattus no
C 31	16.8	84.0	187564	9	CNS0000E	AL049835 Human chr
C 32	16.8	84.0	194140	1	AF242881	AF242881 Agrobacte
C 33	16.8	84.0	203047	10	AL928922	AL928922 Mouse DNA
C 34	16.8	84.0	216397	2	AC119634	AC119634 Rattus no
C 35	16.8	84.0	228273	2	AC111408	AC111408 Rattus no
C 36	16.8	84.0	235971	2	AC124950	AC124950 Rattus no
C 37	16.8	84.0	237857	2	AC093982	AC093982 Rattus no
C 38	16.8	84.0	242569	2	AC122861	AC122861 Mus muscu
C 39	16.8	84.0	246789	2	AC119590	AC119590 Rattus no
C 40	16.8	84.0	251516	2	AC130168	AC130168 Rattus no
C 41	16.8	84.0	252646	2	AC137013	AC137013 Rattus no
C 42	16.8	84.0	294307	2	AC095267	AC095267 Rattus no
C 43	16.4	82.0	488	5	AF292650	AF292650 Osteoglos
C 44	16.4	82.0	512	11	BV004810	BV004810 S208P6428
C 45	16.4	82.0	44557	1	AF305077	AF305077 Anaplasm
C 46	16.4	82.0	69161	2	AC101153	AC101153 Mus muscu
C 47	16.4	82.0	69161	2	AC101153	AC101153 Mus muscu
C 48	16.4	82.0	72887	2	AC017882	AC017882 Drosophi
C 49	16.4	82.0	117482	9	AC105028	AC105028 Homo sapi
C 50	16.4	82.0	124230	10	AC005329	AC005329 Mouse BAC
C 51	16.4	82.0	143180	2	AC120596	AC120596 Rattus no
C 52	16.4	82.0	160931	3	AC093498	AC093498 Drosophi
C 53	16.4	82.0	161841	9	AC135279	AC135279 Homo sapi
C 54	16.4	82.0	162063	9	AC015553	AC015553 Homo sapi
C 55	16.4	82.0	169869	2	AC005867	AC005867 Homo sapi
C 56	16.4	82.0	176121	9	AC131011	AC131011 Homo sapi
C 57	16.4	82.0	176932	9	AL157884	AL157884 Human DNA
C 58	16.4	82.0	181121	10	AC058789	AC058789 Mus muscu
C 59	16.4	82.0	184263	9	AC004805	AC004805 Homo sapi
C 60	16.4	82.0	186451	10	AC122211	AC122211 Mus muscu
C 61	16.4	82.0	193743	2	AC125024	AC125024 Mus muscu
C 62	16.4	82.0	195988	2	AC141493	AC141493 Rattus no
C 63	16.4	82.0	200207	2	AC007867	AC007867 Homo sapi
C 64	16.4	82.0	208030	10	AL672276	AL672276 Mouse DNA
C 65	16.4	82.0	215532	9	AC010883	AC010883 Homo sapi



C 66	16.4	82.0	219239	2	AC123661	AC123661	Mus muscu	C 139	15.8	79.0	1170	6	BD092631	BD092631	Identific
C 67	16.4	82.0	224697	2	AC147234	AC147234	Mus muscu	140	15.8	79.0	1471	3	AV052063	AV052063	Drosophil
C 68	16.4	82.0	234399	2	AC095649	AC095649	Rattus no	141	15.8	79.0	1784	8	AY156979	AY156979	Zea mays
C 69	16.4	82.0	276193	2	AC094506	AC094506	Rattus no	C 142	15.8	79.0	1851	6	AR396960	AR396960	Sequence
C 70	16.4	82.0	303335	3	AE003538	AE003538	Drosophil	C 143	15.8	79.0	4400	6	AR227535	AR227535	Sequence
C 71	16.4	82.0	307120	1	CNSPAX03	AX041919	Pyrococcu	C 144	15.8	79.0	4761	8	SPADAITCP	SPADAITCP	X91498 S.pombe ada
C 72	16.4	82.0	349980	6	AX061938	AX061938	Sequence	C 145	15.8	79.0	8546	6	BD193552	BD193552	Enterococ
C 73	16.4	80.0	258	6	AX060999	AX060999	Sequence	C 146	15.8	79.0	8946	1	AF527793	AF527793	Flavobact
C 74	16.4	80.0	258	6	AX060999	AX060999	Sequence	C 147	15.8	79.0	13695	1	AE000586	AE000586	Helicobac
C 75	16.4	80.0	1394	1	HPY418316	AD042522	Helicobac	C 148	15.8	79.0	14987	1	AE015466	AE015466	Shewanell
C 76	16.4	80.0	1545	1	AY155232	AY155232	Helicobac	C 149	15.8	79.0	16090	9	AL354812	AL354812	Human DNA
C 77	16.4	80.0	195293	9	AL354795	AL354795	Human DNA	C 150	15.8	79.0	21522	1	CEC00811	CEC00811	Salmonell
C 78	15.8	79.0	582	6	AX660988	AX660988	Sequence	C 151	15.8	79.0	26321	3	CEC043107	CEC043107	Z46676 Caenorhabdi
C 79	15.8	79.0	645	11	EV017308	EV017308	Sequence	C 152	15.8	79.0	32752	3	LMFL8806	LMFL8806	AL084603 Caenorhab
C 80	15.8	79.0	648	6	AX790657	AX790657	Sequence	C 153	15.8	79.0	37386	3	NCB7F21	NCB7F21	AL389901 Neurospor
C 81	15.8	79.0	673	6	AX789789	AX789789	Sequence	C 154	15.8	79.0	40915	8	SPAC4G9	SPAC4G9	AF016674 Caenorhab
C 82	15.8	79.0	759	6	AX654628	AX654628	Sequence	C 155	15.8	79.0	41645	3	AF016674	AF016674	AF016674 Caenorhab
C 83	15.8	79.0	1034	6	AX787586	AX787586	Sequence	C 156	15.8	79.0	42544	8	SPSC106	SPSC106	AL110295 S.pombe c
C 84	15.8	79.0	1034	6	AX789421	AX789421	Sequence	C 157	15.8	79.0	45974	9	AL138883	AL138883	Human DNA
C 85	15.8	79.0	1050	5	AF549127	AF549127	Cottus ba	C 158	15.8	79.0	55802	2	AC044912	AC044912	AL044912 Homo sapi
C 86	15.8	79.0	1071	5	AF549120	AF549120	Cottus co	C 159	15.8	79.0	59412	8	NCB7F21	NCB7F21	AL389901 Neurospor
C 87	15.8	79.0	1072	5	AF549119	AF549119	Cottus co	C 160	15.8	79.0	60919	2	AC017150	AC017150	AC017150 Drosophil
C 88	15.8	79.0	1072	5	AF549145	AF549145	Cottus cf	C 161	15.8	79.0	61897	10	AL844423	AL844423	Mouse DNA
C 89	15.8	79.0	1073	5	AF549112	AF549112	Cottus ca	C 162	15.8	79.0	68769	9	AC006027	AC006027	AC006027 Homo sapi
C 90	15.8	79.0	1073	5	AF549113	AF549113	Cottus ca	C 163	15.8	79.0	72327	3	AC004713	AC004713	AC004713 Drosophil
C 91	15.8	79.0	1073	5	AF549115	AF549115	Cottus th	C 164	15.8	79.0	73257	2	AC101211	AC101211	AC101211 Mus muscu
C 92	15.8	79.0	1073	5	AF549113	AF549113	Cottus th	C 165	15.8	79.0	78062	10	AL844572	AL844572	AL844572 Mouse DNA
C 93	15.8	79.0	1079	5	AF549110	AF549110	Cottus ca	C 166	15.8	79.0	87350	2	AC009752	AC009752	AC009752 Homo sapi
C 94	15.8	79.0	1080	5	AF549111	AF549111	Cottus ca	C 167	15.8	79.0	92141	2	AC021565	AC021565	AC021565 Homo sapi
C 95	15.8	79.0	1080	5	AF549114	AF549114	Cottus th	C 168	15.8	79.0	95747	9	AL358175	AL358175	AL358175 Human DNA
C 96	15.8	79.0	1080	5	AF549121	AF549121	Cottus ca	C 169	15.8	79.0	95933	9	AC004874	AC004874	AC004874 Homo sapi
C 97	15.8	79.0	1080	5	AF549112	AF549112	Cottus ca	C 170	15.8	79.0	103313	5	AL954844	AL954844	AL954844 Zebrafish
C 98	15.8	79.0	1080	5	AF549125	AF549125	Cottus ba	C 171	15.8	79.0	103992	2	AC102708	AC102708	AC102708 Mus muscu
C 99	15.8	79.0	1080	5	AF549126	AF549126	Cottus ba	C 172	15.8	79.0	104802	9	AL161736	AL161736	AL161736 Human DNA
C 100	15.8	79.0	1080	5	AF549128	AF549128	Cottus cf	C 173	15.8	79.0	109210	2	AC144760	AC144760	AC144760 Medicago
C 101	15.8	79.0	1080	5	AF549129	AF549129	Cottus cf	C 174	15.8	79.0	110000	2	AC092789_1	AC092789_1	Continuation (2 of
C 102	15.8	79.0	1080	5	AF549131	AF549131	Cottus cf	C 175	15.8	79.0	110492	2	AP004139	AP004139	AP004139 Oryza sat
C 103	15.8	79.0	1080	5	AF549132	AF549132	Cottus cf	C 176	15.8	79.0	112292	9	HSU238D15	HSU238D15	AP096771 Human DNA
C 104	15.8	79.0	1080	5	AF549133	AF549133	Cottus cf	C 177	15.8	79.0	113415	2	AP004860	AP004860	AP004860 Oryza sat
C 105	15.8	79.0	1080	5	AF549134	AF549134	Cottus cf	C 178	15.8	79.0	114964	10	AL691507	AL691507	AL691507 Mouse DNA
C 106	15.8	79.0	1080	5	AF549135	AF549135	Cottus cf	C 179	15.8	79.0	125020	9	AF429315	AF429315	AF429315 Homo sapi
C 107	15.8	79.0	1080	5	AF549136	AF549136	Cottus cf	C 180	15.8	79.0	135330	10	AL772165	AL772165	AL772165 Mouse DNA
C 108	15.8	79.0	1080	5	AF549137	AF549137	Cottus cf								
C 109	15.8	79.0	1080	5	AF549138	AF549138	Cottus cf								
C 110	15.8	79.0	1080	5	AF549139	AF549139	Cottus cf								
C 111	15.8	79.0	1080	5	AF549140	AF549140	Cottus cf								
C 112	15.8	79.0	1080	5	AF549141	AF549141	Cottus cf								
C 113	15.8	79.0	1080	5	AF549144	AF549144	Cottus cf								
C 114	15.8	79.0	1080	5	AF549158	AF549158	Cottus hy								
C 115	15.8	79.0	1080	5	AF549159	AF549159	Cottus hy								
C 116	15.8	79.0	1080	5	AF549160	AF549160	Cottus hy								
C 117	15.8	79.0	1080	5	AF549161	AF549161	Cottus hy								
C 118	15.8	79.0	1140	5	AY050723	AY050723	Caranx ru								
C 119	15.8	79.0	1141	5	AY116342	AY116342	Abyssocott								
C 120	15.8	79.0	1141	5	AY116343	AY116343	Asprocott								
C 121	15.8	79.0	1141	5	AY116344	AY116344	Asprocott								
C 122	15.8	79.0	1141	5	AY116345	AY116345	Asprocott								
C 123	15.8	79.0	1141	5	AY116346	AY116346	Asprocott								
C 124	15.8	79.0	1141	5	AY116347	AY116347	Cottinell								
C 125	15.8	79.0	1141	5	AY116349	AY116349	Limnococt								
C 126	15.8	79.0	1141	5	AY116350	AY116350	Limnococt								
C 127	15.8	79.0	1141	5	AY116351	AY116351	Limnococt								
C 128	15.8	79.0	1141	5	AY116352	AY116352	Limnococt								
C 129	15.8	79.0	1141	5	AY116355	AY116355	Comephoru								
C 130	15.8	79.0	1141	5	AY116356	AY116356	Comephoru								
C 131	15.8	79.0	1141	5	AY116357	AY116357	Batrachoc								
C 132	15.8	79.0	1141	5	AY116358	AY116358	Batrachoc								
C 133	15.8	79.0	1141	5	AY116359	AY116359	Cottocome								
C 134	15.8	79.0	1141	5	AY116360	AY116360	Cottocome								
C 135	15.8	79.0	1141	5	AY116361	AY116361	Cottus ke								
C 136	15.8	79.0	1141	5	AY116362	AY116362	Paracottu								
C 137	15.8	79.0	1141	5	AY116363	AY116363	Cottus ba								
C 138	15.8	79.0	1141	5	AY116367	AY116367	Cottus si								

## ALIGNMENTS

RESULT 1	AP001134/c	158412 bp	DNA	linear	HTG 30-MAY-2000
LOCUS	AP001134	Homo sapiens chromosome 18 clone RP11-785E17 map 18q12	WORKING		
DEFINITION	DRAFT SEQUENCE, 23 unordered pieces.				
ACCESSION	AP001134				
VERSION	AP001134.2	GI:8118426			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 158412)				
TITLE	Hattori.M., Ishii.K., Toyoda.A., Taylor.T.D., Hong-Seog.P., Fujiyama.A., Yada.T., Totoki.Y., Watanabe.H. and Sakaki.Y.				
JOURNAL	Homo sapiens 158,412 genomic DNA of 18q12				
REFERENCE	Published Only in DataBase (2000)				
AUTHORS	2 (bases 1 to 158412)				
TITLE	Hattori.M., Ishii.K., Toyoda.A., Taylor.T.D., Hong-Seog.P., Fujiyama.A., Yada.T., Totoki.Y., Watanabe.H. and Sakaki.Y.				
JOURNAL	Submitted (04-FEB-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (3-mail:hattori@gsc.riken.go.jp).				
	URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,				



## ORIGIN

misc\_feature 144987..147684 /note="assembly\_fragment"  
 misc\_feature 147785..149740 /note="assembly\_fragment"  
 misc\_feature 149841..152517 /note="assembly\_fragment"  
 misc\_feature 152618..154387 /note="assembly\_fragment"  
 misc\_feature 154488..157068 /note="assembly\_fragment"  
 misc\_feature 157169..158412 /note="assembly\_fragment"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 159412;  
 Best Local Similarity 95.0%; Pred. No. 47;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20

Db 74809 TTAAGAGCGCTCCAAAGCC 74790

## RESULT 2

BC034803/c  
 LOCUS BC034803 2144 bp mRNA linear PRI 26-JUL-2002  
 DEFINITION Homo sapiens, clone IMAGE:4821395, mRNA.

ACCESSION BC034803

VERSION BC034803.1 GI:21961560

KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2144)

Strausberg, R.

Direct Submission

Submitted (24-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CNA Library Arrayed by: The I.M.A.G.E.B. Consortium (ILNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E.B. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAC Plate: 71 Row: c Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: Excellent library of

origin.

Location/Qualifiers

1..2144

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4821395"

/tissue\_type="Testis"

/clone\_lib="NIH\_MGC\_97"

/lab\_host="DH10B"

/note="Vector: pBluescript"

## FEATURES

## source

Direct Submission

Unpublished

2 (bases 1 to 132953)

Worley, K.C.

Direct Submission

Submitted (28-FEB-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

Query Match 90.0%; Score 18; DB 9; Length 2144;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20

Db 420 AAAGAGCGCTCCAAAGCC 403

## RESULT 3

AC024221/c

LOCUS AC024221

DEFINITION Homo sapiens 3 BAC RP11-227J5 (Rosenwell Park Cancer Institute Human

BAC Library) complete sequence.

ACCESSION AC024221

VERSION AC024221.23 GI:21205011

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 132953)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oman, F.R., Allen, C.,

Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,

Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.I., Byrd, N.C.,

Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earnhart, C., Edgar, J., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Forgy, J., Foster, P., Prantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., He, X.,

Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,

Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,

Jackson, B.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,

Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,

Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,

Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,

Louise, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,

Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,

Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,

Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,

Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,

Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,

Nickerson, E., Nwokenkwu, S., Ogih, M., Okwuonu, G., Oragunye, N.,

Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,

Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M.,

Rojas, A., Rojucokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S.,

Scott, G., Shen, H., Shooshari, N., Sisson, I., Sodergren, E.,

Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,

Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,

Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmami, K., Vasquez, L.,

Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,

Warren, R., Washington, C., Watlington, S., Williams, G.,

Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,

Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and

Gibbs, R.

```

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (21-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (29-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT
On May 25, 2002 this sequence version replaced gi:20986394.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:13389-1402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.
FEATURES
Source
1. 132953
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-227J5"
misc_feature
1. 2036
/note="overlaps bases 122666..124701 of clone AC079594"
/function="Clone overlap"
repeat_region
complement(35..332)
/rpt_family="AluJb"
repeat_region
995..1028
/rpt_family="AT-rich"
repeat_region
complement(1298..1602)
/rpt_family="AluJb"

```

```

repeat_region
complement(1889..2090)
/rpt_family="AluJb"
repeat_region
2770..2808
/rpt_family="TTTA)n"
repeat_region
complement(2809..3094)
/rpt_family="AluJb"
repeat_region
3313..3372
/rpt_family="MER2"
repeat_region
complement(3379..3435)
/rpt_family="AluJ/FLAM"
repeat_region
3576..3749
/rpt_family="L2"
STS
3648..3796
/standard_name="15013"
STS
3648..3796
/standard_name="69663"
repeat_region
460..4247
/rpt_family="LIP"
repeat_region
4246..9783
/rpt_family="L1PA6"
STS
8111..8390
/standard_name="15245"
repeat_region
complement(10336..10612)
/rpt_family="MIR"
repeat_region
11396..11451
/rpt_family="MER5B"
repeat_region
complement(11458..12087)
/rpt_family="L1VA2"
repeat_region
12834..12866
/rpt_family="(TGA)n"
repeat_region
13142..13161
/rpt_family="(A)n"
repeat_region
14076..14103
/rpt_family="AT-rich"
repeat_region
14383..14427
/rpt_family="AT-rich"
repeat_region
15029..15069
/rpt_family="MIR"
repeat_region
complement(16234..16523)
/rpt_family="AluJb"
STS
17157..17435
/standard_name="134217"
repeat_region
17723..17856
/rpt_family="L1M4"
repeat_region
complement(17858..17989)
/rpt_family="AluJb"
repeat_region
complement(18232..18755)
/rpt_family="MER1A"
repeat_region
18951..19215
/rpt_family="AluJb"
repeat_region
complement(21057..21293)
/rpt_family="L2"
repeat_region
21326..21852
/rpt_family="MLT1F1"
repeat_region
22483..22511
/rpt_family="AT-rich"
repeat_region
23197..24501
/rpt_family="L1MC4"
repeat_region
complement(24573..24624)
/rpt_family="MIR"
repeat_region
24792..24810
/rpt_family="MLT1K"
repeat_region
complement(24811..25083)
/rpt_family="AluSx"
repeat_region
25084..25136
/rpt_family="MLT1K"
repeat_region
25735..26018

```

Query Match 90.0%; Score 18; DB 9; Length 132953;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20

Db 95239 AAAGAGCGCTCCAAAGCC 95222

|||||

RESULT 4

AC018850

LOCUS

DEFINITION

AC018850 Homo sapiens chromosome 3 clone RP11-416D22 map 3, WORKING DRAFT SEQUENCE, 21 unordered pieces.

AC018850 161087 bp DNA linear HTG 26-MAY-2000

AC018850 3 GI:8096873

VERSION

KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 161087)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 3, Clone RP11-416D22

REFERENCE

Unpublished

2 (bases 1 to 161087)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lander,T., Lechoczky,J., Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McEwan,P., McGuck,A., McKernan,K., Meldrum,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliviar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spence,J., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (21-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On May 26, 2000 this sequence version replaced gi:6716059.

All repeats were identified using RepeatMasker:

Smith, A.P.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WISR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L1824

Center clone name: 416.D.22

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 142691 bases at least Q40

Consensus quality: 150851 bases at least Q30

Consensus quality: 154724 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 159087; sum-of-contigs

Quality coverage: 3.3 in Q20 bases; agarose-fp

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

/note="assembly_fragment"
44621..52862
/note="assembly_fragment"
52963..60848
/note="assembly_fragment"
60949..68864
/note="assembly_fragment"
68965..75713
/note="assembly_fragment"
75814..84405
/note="assembly_fragment"
84506..96213
/note="assembly_fragment"
96314..107681
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
107782..133144
/note="assembly_fragment"
133245..161087
/note="assembly_fragment"

```

ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 161087;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 AAAGAGCGTCCAAAGCC 20
Db 72725 AAAGAGCGTCCAAAGCC 72742

AC124946 170067 bp DNA linear HTG 29-JUN-2002
LOCUS Homo sapiens clone Rp11-416D22, *** SEQUENCING IN PROGRESS ***, 3
DEFINITION unordered pieces.
ACCESSION AC124946
VERSION AC124946.3 GI:21539690
KEYWORDS HTG; HTGS PHASE1; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 170067)
Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,
Albrooks S.L., Amarantunge H.C., Are J.R., Ayele M., Banks T.,
Barbaria J., Benton J., Binage K., Blankenburg K., Bonnin D.,
Bouck J., Bowler S., Brieva M., Brown E., Brown M., Bryant N.P.,
Buhay C., Burck P., Burkett C., Burrell K.L., Byrd N.C.,
Carroll T.F., Carter M., Cavazos S.R., Chacko G., Chavez D.,
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,
Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
Gorelli J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,
Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,
Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B.,
Homs F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.B.,
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,
Karlssoon E., Kelly S., Khan U., King L., Korvah J., Kovar C.,
Kratovic J., Kureishi A., Landry N., Leal B., Lewis L.C., Lewis L.,
Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louisseged H.,
Lozano R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,
Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,
Massey E., Mawhiney E., McLeod M.P., Meador M., Mei G., Metzker M.,
Miner G., Miner Z., Mitchell T., Mchabbat K., Morgan M., Morris S.,
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,
Nguyen N., Nickerson E., Nwokenkwo S., Oguh M., Okwuonu G.,

```

```

Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
Rivers M., Rojas A., Rojubokan I., Rolfe M., Ruiz S., Savary G.,
Scherer S., Scott G., Shen H., Shoostari N., Sisson I.,
Sodergren E., Sonaike T., Sparks A., Stanley H., Stone H., Tang H.,
Sutton A., Svatek A., Taber P., Tamerisa A., Tamerisa K., Thomas S.,
Tansey J., Taylor C., Taylor T., Telford B., Thomas R., Thomas S.,
Tumani K., Vasquez L., Vera V., Villalob D., Vinson R., Wang Q.,
Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,
Williams G., Williamson A., Wleczyk R., Wooden S., Worley K.,
Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,
Weinstock G. and Gibbs R.
Direct Submission
Unpublished
2 (bases 1 to 170067)
Worley K.C.
Direct Submission
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 170067)
Worley K.C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 23, 2002 this sequence version replaced gi:21539046.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HEED
Center clone name: Rp11-416D22
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 168450 bases at least Q40
Consensus quality: 168504 bases at least Q30
Consensus quality: 168514 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 15686: contig of 15686 bp in length
* 15687 15786: gap of unknown length
* 15787 61007: contig of 45221 bp in length
* 61008 61107: gap of unknown length
* 61108 170067: contig of 108960 bp in length.
* Location/Qualifiers
* 1..170067
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /clone="Rp11-416D22"

```

FEATURES

source

ORIGIN

```

Query Match 90.0%; Score 18; DB 2; Length 170067;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGTCCAAAGCC 20
|||||
Db 164720 AAAGAGCGTCCAAAGCC 164737

```

## RESULT 6

AC012342/c

LOCUS

DEFINITION

AC012342 Homo sapiens chromosome 3 clone RP11-436F13 map 3, WORKING DRAFT  
SEQUENCE, 22 unordered pieces.

AC012342

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

COMMENT

```

* 3767 3866: gap of 100 bp
* 3867 5900: contig of 2034 bp in length
* 5901 6001: gap of 100 bp
* 6001 7140: contig of 1140 bp in length
* 7141 7240: gap of 100 bp
* 7241 9004: contig of 1764 bp in length
* 9005 9104: gap of 100 bp
* 9105 10107: contig of 1003 bp in length
* 10108 10207: gap of 100 bp
* 10208 12040: contig of 1833 bp in length
* 12041 12141: gap of 100 bp
* 12141 14132: contig of 1992 bp in length
* 14133 14232: gap of 100 bp
* 14233 18334: contig of 4102 bp in length
* 18335 18434: gap of 100 bp
* 18435 23327: contig of 4893 bp in length
* 23328 23427: gap of 100 bp
* 23428 31464: contig of 8037 bp in length
* 31465 31564: gap of 100 bp
* 31565 40161: contig of 8597 bp in length
* 40162 40261: gap of 100 bp
* 40262 48428: contig of 8167 bp in length
* 48429 48528: gap of 100 bp
* 48529 61353: contig of 12825 bp in length
* 61354 72822: contig of 11369 bp in length
* 72823 72922: gap of 100 bp
* 72923 85724: contig of 12802 bp in length
* 85725 85924: gap of 100 bp
* 85925 96778: contig of 10854 bp in length
* 96779 96778: gap of 100 bp
* 96779 112114: contig of 15336 bp in length
* 112115 112114: gap of 100 bp
* 112115 129962: contig of 17748 bp in length
* 129963 130062: gap of 100 bp
* 130063 145946: contig of 15884 bp in length
* 145947 188526: contig of 42480 bp in length.
* 146047 Location/Qualifiers
      1. .188526
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="3"
        /map="3"
        /clone="RP11-436F13"
        /clone_lib="RP11-11 Human Male BAC"
      misc_feature
      1. .1030
        /note="assembly_fragment"
      misc_feature
      131. .2538
        /note="assembly_fragment"
      misc_feature
      2639. .3766
        /note="assembly_fragment"
      misc_feature
      3867. .5900
        /note="assembly_fragment"
      misc_feature
      6001. .7140
        /note="assembly_fragment"
      misc_feature
      7241. .9004
        /note="assembly_fragment"
      misc_feature
      9105. .10107
        /note="assembly_fragment"
      misc_feature
      10208. .12040
        /note="assembly_fragment"
      misc_feature
      12141. .14132
        /note="assembly_fragment"
      misc_feature
      14233. .18334
        /note="assembly_fragment"
      misc_feature
      18435. .23327
        /note="assembly_fragment"
      misc_feature
      23428. .31464
        /note="assembly_fragment"
      misc_feature
      31565. .40161
        /note="assembly_fragment"

```

## FEATURES

source

```

----- Genomic Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1812
Center clone name: 436 F.13
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152028 bases at least Q40
Consensus quality: 172958 bases at least Q30
Consensus quality: 181270 bases at least Q20
Insert size: 182000; agarose-fp
Quality coverage: 4.8 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1030: contig of 1030 bp in length
* 1031 1130: gap of 100 bp
* 1131 2538: contig of 1408 bp in length
* 2539 2639: gap of 100 bp
* 2639 3766: contig of 1128 bp in length

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 128.353 Seconds

(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819B-36

Perfect score: 20

Sequence: 1 ttaagagcgctccaaagcc 20

Scoring table: IDENTITY\_NJC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

1: Geneseqn29Jan04.\*

2: Geneseqn1980s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	6	AAL48182 Human HLA
2	20	100.0	20	6	AAL48202 Human HLA
3	20	100.0	20	6	AAL48200 Human HLA
4	20	100.0	20	6	AAL48166 Human HLA
5	18	90.0	545	5	AAS64515 DNA encod
6	18	90.0	2076	5	AAS64519 DNA encod
7	16.4	82.0	349980	5	AAL41224 Pyrococcus
8	16	80.0	258	3	AAC18777 Human sec
9	16	80.0	1800	2	AAC18777 Human sec
10	15.8	79.0	582	8	ADA49275 Maize gen
11	15.8	79.0	648	6	ABX66962 Helicobac
12	15.8	79.0	673	6	ABX66528 Helicobac
13	15.8	79.0	759	7	ADA71175 Rice gene
14	15.8	79.0	1024	6	ABX66344 Helicobac
15	15.8	79.0	1024	6	ABX65416 Helicobac
16	15.8	79.0	1061	4	ABL02097 Drosophil
17	15.8	79.0	1110	4	AAS3627 Helicobac
18	15.8	79.0	1170	2	ACA34729 Prokaryot
19	15.8	79.0	1170	2	AAX14331 H. pylori
20	15.8	79.0	1581	7	ACA51800 Prokaryot
21	15.8	79.0	3581	4	ABL02096 Drosophil
22	15.8	79.0	8546	2	AAX13083 Enterococ
23	15.8	79.0	8546	6	ABS98878 Enterococ

24	15.8	79.0	349980	5	AAH68528	Aab68528 C glutami
25	15.4	77.0	224	5	ABV19964	Abv19964 Human pro
26	15.4	77.0	236	5	ABV49725	Abv49725 Human pro
27	15.4	77.0	1176	6	AD339110	Ad339110 Human lun
28	15.4	77.0	1203	4	ABL26151	Ab126151 Drosophil
29	15.4	77.0	1440	2	AAX14179	Aax14179 H. pylori
30	15.4	77.0	3402	4	ABL10673	Ab110673 Drosophil
31	15.4	77.0	5583	4	ABL10672	Ab110672 Drosophil
32	15.4	77.0	10766	4	AAK90312	Aak90312 Human dig
33	15.4	77.0	10766	5	AAS39895	Aas39895 Genomic s
34	15.4	77.0	10766	8	ADB32855	Adb32855 Human nov
35	15.2	76.0	46	2	AAK26508	Aax26508 WO 990919
36	15.2	76.0	187	6	ABL79993	Ab179993 Human ova
37	15.2	76.0	414	8	ACH18032	ACH18032 Human adu
38	15.2	76.0	422	5	AAS66006	Aas66006 DNA encod
39	15.2	76.0	627	2	AAQ87825	Aaq87825 Agmenellu
40	15.2	76.0	747	5	AAH67991	Aah67991 C glutami
41	15.2	76.0	762	6	ABK74665	Abk74665 Bacillus
42	15.2	76.0	765	2	AAV24610	Aav24610 H. pylori
43	15.2	76.0	768	2	AAV24886	Aav24886 H. pylori
44	15.2	76.0	768	2	AAV75824	Aax75824 H. pylori
45	15.2	76.0	1110	4	AAS53795	Aas53795 Helicobac
46	15.2	76.0	1209	7	ADA70113	Ada70113 Rice gene
47	15.2	76.0	1686	8	ADB09055	Adb09055 Alloiococ
48	15.2	76.0	1686	8	ADB09057	Adb09057 Alloiococ
49	15.2	76.0	1686	8	ADB09059	Adb09059 Alloiococ
50	15.2	76.0	1727	2	AAQ11036	Aaq11036 Gene encod
51	15.2	76.0	1838	7	ACD19454	Acd19454 cDNA encod
52	15.2	76.0	2044	1	AAN81505	Aan81505 DNA encod
53	15.2	76.0	2061	4	AAS53803	Aas53803 Helicobac
54	15.2	76.0	2136	2	AAQ26642	Aaq26642 ISRml in
55	15.2	76.0	2253	5	AAH67990	Aah67990 C glutami
56	15.2	76.0	2466	7	ACA00124	Aca00124 C. glutam
57	15.2	76.0	2586	9	ADD13666	Adh13666 C. glutam
58	15.2	76.0	2884	4	AAH14544	Aah14544 Human cDN
59	15.2	76.0	3159	9	ADL45409	Adl45409 Rat gene
60	15.2	76.0	3240	4	ABL02064	Ab102064 Drosophil
61	15.2	76.0	3241	4	ABL20820	Ab120820 Drosophil
62	15.2	76.0	3241	4	ABL20822	Ab120822 Drosophil
63	15.2	76.0	3411	2	AAT38950	Aat28950 Helicobac
64	15.2	76.0	4018	7	ABZ34827	Abz34827 Coding se
65	15.2	76.0	4234	2	AAQ87819	Aaq87819 Agmenellu
66	15.2	76.0	8396	3	AAS59353	Aaz59353 Human STP
67	15.2	76.0	8672	9	ADD18777	Adh18777 Human dis
68	15.2	76.0	34503	4	AAK79621	Aak79621 Human imm
69	15.2	76.0	3503	4	AAK67032	Aak67032 Human imm
70	15.2	76.0	35100	2	AAV73803	Aav73803 KSHV LTR
71	15.2	76.0	48551	6	AAS20800	Aas20800 Clostridi
72	15.2	76.0	110000	8	ADB12064	Continuation (8 of
73	15.2	76.0	137507	2	AAV19941	Aav19941 KSHV long
74	15.2	76.0	349980	5	AAH68533	Aah68533 C glutami
75	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
76	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
77	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
78	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
79	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
80	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
81	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
82	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
83	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
84	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
85	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
86	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
87	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
88	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
89	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
90	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
91	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
92	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
93	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
94	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
95	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu
96	15.2	76.0	349980	5	AAF86431	Aaf86431 Pyrococu



	C	97	14.8	74.0	408	6	Abs17785	Human	Sen		c	170	14.4	72.0	213	4	ABA36125	Probe #14
	98	14.8	74.0	480	5	AEA06274	Soy bean				C 171	14.4	72.0	213	4	AAK43307	Human bor	
	99	14.8	74.0	480	5	Aba06273	Soy bean				C 172	14.4	72.0	213	4	Rak17500	Human bra	
	100	14.8	74.0	481	9	ADB51204	Primary r				C 173	14.4	72.0	213	4	Abs42934	Human liv	
	101	14.8	74.0	534	4	AAI17148	Probe #70				C 174	14.4	72.0	213	5	Aai09659	Probe #96	
	102	14.8	74.0	534	4	ABa61630	Human foe				C 175	14.4	72.0	213	6	Abs17389	Human gen	
	103	14.8	74.0	534	4	AAI41542	Probe #10				C 176	14.4	72.0	231	9	Add24495	DNA polym	
	104	14.8	74.0	534	4	AbA29292	Probe #77				C 177	14.4	72.0	231	9	AGd24494	DNA polym	
	105	14.8	74.0	534	4	AAK35825	Human bon				C 178	14.4	72.0	351	3	AAC04990	Human sec	
	106	14.8	74.0	534	4	AAK09930	Human bra				C 179	14.4	72.0	368	5	AAF64987	Novel hum	
	107	14.8	74.0	534	4	AAK09930	Human bra				C 180	14.4	72.0	426	4	RAI14891	Novel #48	
	108	14.8	74.0	534	4	AAK09930	Human bra				C 181	14.4	72.0	426	4	RAI14891	Novel #48	

## ALIGNMENTS

C	112	14.8	74.0	777	6	AAC67928	Listeria
C	113	14.8	74.0	868		AAC47508	Arabidops
C	114	14.8	74.0	870	3	AAC34139	Arabidops
C	115	14.8	74.0	980	5	AAS83196	DNA encod
C	116	14.8	74.0	1050	4	AAL71772	ActA prot
C	117	14.8	74.0	1128	4	AAL71771	ActA prot
C	118	14.8	74.0	1156	6	ABK63440	Rat seque
C	119	14.8	74.0	1156	9	ADB57852	Toxicity-
C	120	14.8	74.0	1156	9	ADB52358	Primary r
C	121	14.8	74.0	1156	9	ADB85151	Rat UDP- G
C	122	14.8	74.0	1229	4	ABL05839	Drosophil
C	123	14.8	74.0	1467	5	AAR55041	C glutami KW

C 125	14.8	74.0	1836	AAH14772	Human cDN	OS
C 126	14.8	74.0	1841	AAH14772	DNA encod	XX
C 127	14.8	74.0	1878	AA573554	DNA encod	XX
C 128	14.8	74.0	1920	AAV37026	Listeria	PK
C 129	14.8	74.0	1920	ABQ69832	Listeria	XX
C 130	14.8	74.0	1920	ABQ69832	Listeria	XX
C 131	14.8	74.0	1920	ABQ67930	Listeria	PD
C 132	14.8	74.0	2016	AA573555	DNA encod	XX
C 133	14.8	74.0	2631	ACA53272	Prokaryot	PF
C 134	14.8	74.0	2726	AA526593	Human cDN	XX
C 135	14.8	74.0	2726	ABX73934	Human nov	PR
C 136	14.8	74.0	2744	AA526178	Human cDN	XX
C 137	14.8	74.0	2744	ABX73519	Human nov	PA
C 138	14.8	74.0	2850	AAV64278	Human EGR	XX
C 139	14.8	74.0	3074	ABZ82508	Human sec	PI
C 140	14.8	74.0	3293	ACA31588	Prokaryot	XX
C 141	14.8	74.0	3444	ABL05838	Drosophil	DR
C 142	14.8	74.0	3587	ACA35695	Prokaryot	XX
C 143	14.8	74.0	3885	ABL23868	Drosophil	PT
C 144	14.8	74.0	4272	ACA35793	Prokaryot	PT
C 145	14.8	74.0	4272	ABT11094	Human bre	XX
	14.8	74.0	4272	ABX10965	cDNA enco	XX

147	-4.8	74.0	4339	7	ABX62934	Human act	XX
148	-4.8	74.0	4724	2	AAX20542	Polynucle	CC
c 149	-4.8	74.0	6714	4	ABL16321	Drosophil	CC
150	-4.8	74.0	7272	9	ADB56211	Toxicity-	CC
151	-4.8	74.0	7471	3	AAX88789	B. subtil	CC
c 152	-4.8	74.0	11838	4	ABL16320	Drosophil	CC
153	-4.8	74.0	110000	6	ABA03041_02	Continuat	CC
c 154	-4.8	74.0	145831	6	ABL69213	Prostate	CC
c 155	-4.8	74.0	145831	6	ABL66806	Lung can	CC
c 156	-4.8	74.0	145831	6	ABL68588	Kidney ca	CC
c 157	-4.8	74.0	145831	6	ABL62309	Colon ade	CC

C	159		74.0	34980	5	AAb64966	C glutami	xx
	160	14.4	72.0	126	4	ABa69371	Human foe	SQ Sequence 20 BP; 7 A; 6 C; 4 G; 3 T; 0 U; 0 Other;
	161	14.4	72.0	126	4	ABa38307	Probe #14	
	162	14.4	72.0	126	4	ABa43096	Human liv	Query Match 100.0%; Score 20; DB 6; Length 20;
	163	14.4	72.0	126	6	ABS17584	Hunan Sen	Best Local Similarity 100.0%; Pred. No. 0.92;
	164	14.4	72.0	126	6	ABS17584	Hunan Sen	Best Conservative 100.0%; Mismatches 0; Indels 0;

[illegible]

RESULT 2  
AAL48202/c  
ID AAL48202 standard; DNA; 20 BP.  
XX  
AC AAL48202;  
XX  
DT 01-OCT-2002 (first entry)  
XX  
DE Human HLA DPB1 locus polymorphism address tag sequence #2.  
XX  
KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
KW Flow cytometry; human; DPB1; address tag; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200194639-A1.  
XX  
PD 13-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US018590.  
XX  
PR 08-JUN-2000; 2000US-0210759P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI White PS, Torney DC;  
XX  
PW 2002-566450/60.  
XX  
PS Identifying sequences useful as address/capture tags for flow cytometry  
PT based minisequencing, by generating tag sequences and rejecting sequences  
PT based on certain parameters e.g. sequences which form stable hairpins.  
XX  
PS Disclosure; Page 14; 35pp; English.  
XX  
CC The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is an address tag described in the  
CC exemplification of the invention  
XX  
SQ Sequence 20 BP; 3 A; 4 C; 6 G; 7 T; 0 U; 0 Other;  
PS  
XX  
XX The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is an address tag described in the  
CC exemplification of the invention  
XX  
SQ Sequence 20 BP; 3 A; 4 C; 6 G; 7 T; 0 U; 0 Other;  
PS  
XX  
XX Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 TTAAGAGCGCTCCAAAGCC 20  
DB 20 TTAAGAGCGCTCCAAAGCC 1  
XX  
RESULT 3  
AAL48200  
ID AAL48200 standard; DNA; 20 BP.  
XX  
AC AAL48200;  
XX  
DT 01-OCT-2002 (first entry)  
XX  
DE Human HLA DPB1 locus polymorphism multiplex capture sequence #2.  
XX  
KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
KW Flow cytometry; human; DPB1; capture tag; ss.  
XX  
OS Homo sapiens.

XX  
PN WO200194639-A1.  
XX  
PD 13-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US018590.  
XX  
PR 08-JUN-2000; 2000US-0210759P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI White PS, Torney DC;  
XX  
PW 2002-566450/60.  
XX  
PS Identifying sequences useful as address/capture tags for flow cytometry  
PT based minisequencing, by generating tag sequences and rejecting sequences  
PT based on certain parameters e.g. sequences which form stable hairpins.  
XX  
PS Disclosure; Page 14; 35pp; English.  
XX  
CC The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is a capture tag described in the  
CC exemplification of the invention  
XX  
SQ Sequence 20 BP; 7 A; 6 C; 4 G; 3 T; 0 U; 0 Other;  
PS  
XX  
XX Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 TTAAGAGCGCTCCAAAGCC 20  
DB 1 TTAAGAGCGCTCCAAAGCC 20  
XX  
RESULT 4  
AAL48166/c  
ID AAL48166 standard; DNA; 20 BP.  
XX  
AC AAL48166;  
XX  
DT 01-OCT-2002 (first entry)  
XX  
DE Human HLA DPB1 locus polymorphism address tag sequence #2.  
XX  
KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
KW Flow cytometry; human; DPB1; capture tag; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200194639-A1.  
XX  
PD 13-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US018590.  
XX  
PR 08-JUN-2000; 2000US-0210759P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI White PS, Torney DC;  
XX  
PW 2002-566450/60.  
XX

PT Identifying sequences useful as address/capture tags for flow cytometry  
PT based mini-sequencing, by generating tag sequences and rejecting sequences  
PT based on certain parameters e.g. sequences which form stable hairpins.

PS Disclosure; Page 9; 35pp; English.

XX The present invention relates to a method of identifying sequences useful  
CC as address/capture tags, involving rejecting sequences having common sub-  
CC sequences with a sub-sequence length greater than specified number of  
CC bases, and sequences which can form stable hairpins and stable dimers  
CC from a sample of oligonucleotides, and selecting those sequences in the  
CC sample that would hybridise to their respective complements with a high  
CC degree of specificity. The method is useful for identifying a set of  
CC sequences useful as address/capture tags which can be used for  
CC multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
CC cytometry assay. The present sequence is a capture tag described in the  
CC exemplification of the invention

XX Sequence 20 BP; 3 A; 4 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20  
|||  
DB 20 TTAAGAGCGCTCCAAAGCC 1

## RESULT 5

AAS64515/c

ID AAS64515 standard; cDNA; 545 BP.

XX

AC AAS64515;

XX

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #319.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-C0540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

PI WPI; 2001-639362/73.

XX

DR P-PSDB; ABG00328.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

PS Claim 1; SEQ ID NO 319; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC

CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 545 BP; 86 A; 164 C; 171 G; 124 T; 0 U; 0 Other;

Query Match 90.0%; Score 18; DB 5; Length 545;

Best Local Similarity 100.0%; Pred. No. 16;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20

|||  
DB 421 AAAGAGCGCTCCAAAGCC 404

## RESULT 6

AAS64519/c

ID AAS64519 standard; cDNA; 2076 BP.

XX

AC AAS64519;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #323.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

PI WPI; 2001-639362/73.

XX

DR P-PSDB; ABG00332.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

PS Claim 1; SEQ ID NO 323; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC

polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (III). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 2076 BP; 465 A; 603 C; 596 G; 412 T; 0 U; 0 Other;

Query Match 90.0%; Score 18; DB 5; Length 2076;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
DB 555 AAAGAGCGCTCCAAAGCC 538  
|||||

RESULT 7  
AAH41224  
ID AAH41224 standard; DNA: 349980 BP.  
XX  
AC AAH41224;  
XX  
DT 29-OCT-2001 (first entry)  
XX  
DE Pyrococcus abyssi genomic fragment #3.  
XX  
KW Hyperthermophilic archaeon; hyperthermophilic protein; ds.  
XX  
OS Pyrococcus abyssi.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1..49980  
FT /tag= a  
FT /note= "This sequence overlaps with the 3' end of  
AAH41223"  
FT PT 300001..349980  
FT /tag= b  
FT /note= "This sequence overlaps with the 5' end of  
AAH41225"  
FT  
XX PR2792651-A1.  
XX  
PD 27-OCT-2000.  
XX  
XX 21-APR-1999; 99FR-00005034.  
XX  
XX 21-APR-1999; 99FR-00005034.  
XX  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX (IFRE-) IFREMER INST FR RECH EXPL MER.  
XX  
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompre O;  
XX Querellou J, Weissenbach J, Saurin W, Heilig R;  
XX WPI; 2001-126236/14.  
XX  
XX New nucleotide sequences isolated from Pyrococcus abyssi encode proteins useful in industry.  
XX  
XX Claim 1; Page 347-443; 1657pp; French.  
XX  
XX The present invention relates to the genomic sequence of Pyrococcus abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal

vents. The present sequence is a fragment of the genomic sequence of P. abyssi. The 5' end of this sequence overlaps with the 3' end of AAH41223 and the 3' end of this sequence overlaps with the 5' end of AAH41225. The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to 110 degrees centigrade. Note: This patent is in the same patent family as WO2000065062, which contains additional sequences as shown in AAB99132-AAB99143, AAH75903-AAH75920 and AAG66436

Sequence 349980 BP; 94090 A; 78692 C; 78319 G; 98879 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 5; Length 349980;  
Best Local Similarity 94.4%; Pred. No. 2.6e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19  
DB 115436 TAGAGAGCGCTCCAAAGC 115453  
|||||

RESULT 8  
AAC18777/c  
ID AAC18777 standard; cDNA; 258 BP.  
XX  
AC AAC18777;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 22852.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.  
XX  
XX Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
XX (GEST ) GENSET.  
XX  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX Claim 1; SEQ ID NO 22852; 71pp + Sequence Listing; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

Sequence 258 BP; 68 A; 46 C; 65 G; 75 T; 0 U; 4 Other;

Query Match 80.0%; Score 16; DB 3; Length 258;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20  
DB 251 TTAAGAGMCCYTCAAAGCC 232

RESULT 9  
AAV27582  
ID AAV27582 standard; DNA; 1800 BP.  
XX  
AC AAV27582;  
XX  
XX 17-OCT-2003 (revised)  
DT 12-OCT-1998 (first entry)  
XX  
XX Helicobacter pylori flaB gene.  
DE  
XX  
XX Flagellin; flaB gene; vaccine; immunogen; gastritis; peptic ulcer; ss.  
KW  
XX  
OS Helicobacter pylori; strain CCUG 17874.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 138..1682  
FT /\*tag= a  
XX  
XX WO9823288-A1.  
XX  
XX 04-JUN-1998.  
XX  
XX 18-NOV-1997; 97WO-SE001928.  
XX  
XX 25-NOV-1996; 96SE-00004322.  
XX  
XX (ASTR ) ASTRA AB.  
XX  
XX Boelin I, Berglinth T, Mellgard B, Svennerholm A;  
XX  
XX WPI; 1998-322460/28.  
XX P-PSDB; AAW61270.  
XX  
XX Using Helicobacter pylori flagellin polypeptide as immunogen in vaccines  
XX - for treatment or prevention of Helicobacter pylori infection, provide  
XX strong and consistent immune response.  
XX  
XX Disclosure; Page 26-29; 45pp; English.  
XX  
XX This genomic DNA comprises the flaB gene of Helicobacter pylori CCUG  
XX 17874 that codes for a flagellin polypeptide (see AAW61270). It was  
XX cloned from a H. pylori CCUG 17874 genomic library using probes obtained  
XX by PCR amplification of the 5' and 3' regions of the gene. The flagellin  
XX polypeptide FlaB or FlaA (see AAW61269), or its modified but  
XX antigenically equivalent forms, can be used to induce a protective immune  
XX response against H. pylori infection. Also claimed are vaccines  
XX containing FlaA or FlaB and a carrier or diluent. The vaccines are used  
XX to treat or prevent H. pylori infection, particularly in humans.  
XX Flagellin is a strong and consistent antigen that stimulates a local  
XX immune response which decreases or eliminates colonisation of the gastric  
XX mucosa. (Updated on 17-OCT-2003 to standard:se OS field)  
XX  
XX Sequence 1800 BP; 545 A; 347 C; 448 G; 460 T; 0 U; 0 Other;  
XX

Query Match 80.0%; Score 16; DB 2; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 2.1e-02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAGAGCGCTCCAAAGC 19  
DB 464 AAGAGCGCTCCAAAGC 479

RESULT 10

ADA49275  
ID ADA49275 standard; DNA; 582 BP.  
XX  
AC ADA49275;  
XX  
XX 20-NOV-2003 (first entry)  
DT  
XX  
XX Maize gene conferring disease resistance in plants.  
DE  
XX  
XX disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;  
KW maize.  
XX  
XX Zea mays.  
XX  
XX WO2003000906-A2.  
XX  
XX 03-JAN-2003.  
XX  
XX 21-JUN-2002; 2002WO-IB002453.  
XX  
XX 22-JUN-2001; 2001US-0300112P.  
XX 26-SEP-2001; 2001US-035227P.  
XX 22-MAR-2002; 2002US-036653P.  
XX  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;  
XX Katagiri F, Krepes J, Provart N, Ricke D, Zhu T;  
XX WPI; 2003-184052/18.  
XX  
XX New polynucleotide comprising a plant nucleotide sequence having an open  
XX reading frame that encodes a polypeptide associated with disease  
XX resistance, useful for conferring resistance or tolerance to a plant  
XX pathogen.  
XX  
XX Disclosure; SEQ ID NO 1345; 299pp; English.  
XX  
XX The invention relates to a novel isolated polynucleotide comprising a  
XX plant nucleotide sequence having an open reading frame that encodes a  
XX polypeptide associated with disease resistance or its fragment having  
XX substantially the same activity as the full-length polypeptide. The  
XX polynucleotide of the invention is useful for conferring resistance or  
XX tolerance to a plant pathogen. The present sequence represents a gene  
XX conferring disease resistance used in the invention.  
XX  
XX Sequence 582 BP; 129 A; 141 C; 157 G; 155 T; 0 U; 0 Other;  
XX

Query Match 79.0%; Score 15.8; DB 8; Length 582;  
Best Local Similarity 89.5%; Pred. No. 2.3e-02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20  
DB 512 TGAAGAGCGATCCAAAGCC 530

RESULT 11  
ABX66962/c  
ID ABX66962 standard; DNA; 645 BP.  
XX  
XX ABX66962;  
XX  
XX 07-MAY-2003 (first entry)  
DT  
XX  
XX Helicobacter pylori selected interacting domain (SID) DNA #1561.  
DE  
XX  
XX Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX gene; ds.  
XX  
XX Helicobacter pylori.  
XX  
XX WO200266501-A2.  
XX



SQ Sequence 759 BP; 194 A; 183 C; 199 G; 183 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 7; Length 759;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAAGAGCGCTCCAAAGC 19  
DB 714 TTCAAGTGGCTCCAAAGC 696  
RESULT 14  
ABX66344/c  
ID ABX66344 standard; DNA; 1024 BP.  
XX AC  
XX AC  
XX AC  
XX 07-MAY-2003 (first entry)  
XX Helicobacter pylori selected interacting domain (SID) DNA #943.  
XX Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX gene; ds.  
XX Helicobacter pylori.  
XX WO200266501-A2.  
XX 29-AUG-2002.  
XX 28-DEC-2001; 2001WO-EP015428.  
XX 02-JAN-2001; 2001US-0259302P.  
XX (HYBR-) HYBRIGENICS.  
XX (INSP) INST PASTEUR.  
XX Legrain P, Rain J, Collard F, De Reuse H, Labigne A;  
XX WPI; 2002-674910/72.  
XX P-PSDB; ABU51600.  
XX New complexes of protein-protein interactions in Helicobacter pylori,  
XX useful for identifying modulating compounds for treating or preventing  
XX ulcers in mammals.  
XX Claim 7; Page 314; 642pp; English.  
XX The invention describes a complex of protein-protein interactions in  
XX Helicobacter pylori selected from 421 complexes given in the  
XX specification. The complex of protein-protein interactions are useful for  
XX screening for agents which modulate the interaction of proteins.  
XX Modulating compounds which binds to a targeted bacterial protein may be  
XX used for treating or preventing ulcers in a human or animal. This  
XX sequence encodes a selected interacting domain (SID), identified via  
XX protein-protein interactions  
XX Sequence 1024 BP; 337 A; 179 C; 218 G; 290 T; 0 U; 0 Other;  
SQ Query Match 79.0%; Score 15.8; DB 6; Length 1024;  
Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 TTAAGAGCGCTCCAAAGC 20  
DB 77 TATAGAGCGCTCCAAAGC 59  
RESULT 15  
ABX65416/c  
ID ABX65416 standard; DNA; 1024 BP.  
XX AC  
XX ABX65416;

XX 07-MAY-2003 (first entry)  
XX Helicobacter pylori selected interacting domain (SID) DNA #15.  
XX Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX gene; ds.  
XX Helicobacter pylori.  
XX WO200266501-A2.  
XX 29-AUG-2002.  
XX 28-DEC-2001; 2001WO-EP015428.  
XX 02-JAN-2001; 2001US-0259302P.  
XX (HYBR-) HYBRIGENICS.  
XX (INSP) INST PASTEUR.  
XX Legrain P, Rain J, Collard F, De Reuse H, Labigne A;  
XX WPI; 2002-674910/72.  
XX P-PSDB; ABU50672.  
XX New complexes of protein-protein interactions in Helicobacter pylori,  
XX useful for identifying modulating compounds for treating or preventing  
XX ulcers in mammals.  
XX Claim 7; Page 87; 642pp; English.  
XX The invention describes a complex of protein-protein interactions in  
XX Helicobacter pylori selected from 421 complexes given in the  
XX specification. The complex of protein-protein interactions are useful for  
XX screening for agents which modulate the interaction of proteins.  
XX Modulating compounds which binds to a targeted bacterial protein may be  
XX used for treating or preventing ulcers in a human or animal. This  
XX sequence encodes a selected interacting domain (SID), identified via  
XX protein-protein interactions  
XX Sequence 1024 BP; 317 A; 206 C; 234 G; 267 T; 0 U; 0 Other;  
SQ Query Match 79.0%; Score 15.8; DB 6; Length 1024;  
Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 TAAAGAGCGCTCCAAAGC 20  
DB 767 TATAGAGCGCTCCAAAGC 749  
RESULT 16  
ABU02097  
ID ABL02097 standard; cDNA; 1061 BP.  
XX AC  
XX ABL02097;  
XX 26-MAR-2002 (first entry)  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 773.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ss.  
XX Drosophila melanogaster.  
XX WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX

PR 23-MAR-2000; 2000US-019-637P.  
 PR 11-JUL-2000; 2000US-006-4150.  
 XX (PEKE ) PE CORP NY.  
 PA  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 XX WPI; 2001-556860/75.  
 DR P-PSDB; ABB57994.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 XX Claim 1; SEQ ID NO 773; 21pp + Sequence Listing; English.  
 PS  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 XX Sequence 1061 BP; 246 A; 281 C; 263 G; 271 T; 0 U; 0 Other;  
 SQ  
 Query Match 79.0%; Score 15.8; DB 4; Length 1061;  
 Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 TAAAGAGCGCTCCAAAGCC 20  
 DB 138 TAAAGAGCGCCCAAAACC 156  
 RESULT 17  
 AAS53627/c  
 ID AAS53627 standard; DNA; 1110 BP.  
 XX  
 XX AAS53627;  
 AC  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE  
 DE Helicobacter pylori DNA for cellular proliferation protein #81.  
 XX  
 XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
 KW antibacterial; drug design.  
 XX  
 XX Helicobacter pylori.  
 OS  
 XX WO200170955-A2.  
 PN  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 21-MAR-2001; 2001WO-US009180.  
 PF  
 XX 21-MAR-2000; 2000US-0191078P.  
 XX 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207927P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX  
 XX (BLIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 PI  
 PI WPI; 2001-611495/70.  
 XX  
 XX

DR 2-PSDB; AAU35768.  
 XX  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 PT  
 XX  
 XX Claim 27; SEQ ID NO 7264; 511pp; English.  
 PS  
 XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes,  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence encodes an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 XX Sequence 1110 BP; 344 A; 217 C; 253 G; 236 T; 0 U; 0 Other;  
 SQ  
 Query Match 79.0%; Score 15.8; DB 4; Length 1110;  
 Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 TAAAGAGCGCTCCAAAGCC 20  
 DB 772 TATAGAGCGCTTCAAAGCC 754  
 RESULT 18  
 ACA34729/c  
 ID ACA34729 standard; DNA; 1110 BP.  
 XX  
 XX ACA34729;  
 AC  
 XX  
 DT 19-JUN-2003 (first entry)  
 DE  
 DE Prokaryotic essential gene #16386.  
 XX  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX  
 XX Helicobacter pylori.  
 OS  
 XX WO200277183-A2.  
 PN  
 XX  
 XX 03-OCT-2002.  
 PD  
 XX  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (BLIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 PI  
 PI WPI; 2003-029926/02.  
 DR P-PSDB; ABU30859.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening



PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 14; SEQ ID NO 22599; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 1110 BP; 344 A; 217 C; 253 G; 296 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DS 7; Length 1110;  
 Best Local Similarity 89.5%; Pred. No. 2.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 TAAAGAGCGCTCCAAAGCC 20  
 |||||  
 Db 772 TATAGAGCGCTCCAAAGCC 754

RESULT 19  
 AAX14331/c  
 ID AAX14331 standard; DNA; 1170 BP.

XX AAX14331;

XX 31-MAR-1999 (first entry)

XX H. pylori GHPO 202 gene.

XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease; ss.

XX Helicobacter pylori.

OS Key Location/Qualifiers

XX 24..1133

XX /\*tag= a

XX WO9843478-A1.

XX 08-OCT-1998.

XX 01-APR-1998; 98WO-US006371.

PR 01-APR-1997; 97US-00833457.  
 PR 24-JUN-1997; 97US-00881227.  
 PR 29-JUL-1997; 97US-00902615.  
 XX  
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 PA (HUYA-) HUMAN GENOME SCI INC.

PI Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;  
 XX WPI; 1998-542293/46.  
 DR P-FSDS; AAU98612.

XX New isolated Helicobacter polynucleotides - used to develop products for  
 PT the diagnosis, prevention and treatment of Helicobacter infections and  
 PT gastrointestinal diseases.

PS Claim 1; Page 1235-1237; 2054pp; English.

XX This sequence represents a polynucleotide of the invention. It was  
 CC isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.  
 CC The polypeptides can be used for preventing or treating Helicobacter  
 CC infections, and gastroduodenal diseases associated with these infections,  
 CC including acute, chronic, and atrophic gastritis, and peptic ulcer  
 CC diseases, e.g. gastric and duodenal ulcers. They can also be used for the  
 CC production of antibodies. The products can also be used for detection and  
 CC diagnosis

XX Sequence 1170 BP; 368 A; 224 C; 259 G; 319 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DS 2; Length 1170;  
 Best Local Similarity 89.5%; Pred. No. 2.6e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 TAAAGAGCGCTCCAAAGCC 20  
 |||||  
 Db 795 TATAGAGCGCTCCAAAGCC 777

RESULT 20  
 ACA51800  
 ID ACA51800 standard; DNA; 1581 BP.

XX ACA51800;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #33457.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.

OS Salmomella typhi.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-FSDS; ABU47930.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 ; Search time 27.1765 Seconds  
(without alignments)  
408.405 Million cell updates/sec

Title: US-09-877-819b-36

Perfect score: 20

Sequence: 1 ttaagagcgtcccaagcc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCRUS COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	15.8	79.0	1851	4	US-09-134-000C-2975
C 2	15.8	79.0	4400	4	US-09-221-017B-995
C 3	15.2	76.0	1476	4	US-09-489-039A-792
C 4	15.2	76.0	3184	4	US-09-976-594-291
C 5	15.2	76.0	3411	2	US-08-849-480A-1
C 6	15.2	76.0	7152	3	US-09-157-681-29
C 7	15.2	76.0	8396	4	US-09-328-274A-1
C 8	15.2	76.0	8409	3	US-09-157-681-37
C 9	15.2	76.0	35100	3	US-08-770-379-18
C 10	15.2	76.0	35100	3	US-08-757-669A-18
C 11	15.2	76.0	35100	4	US-09-230-371A-18
C 12	15.2	76.0	35100	4	US-09-523-856-36
C 13	15.2	76.0	105	4	US-09-523-856-33
C 14	15.2	76.0	4404	4	US-09-523-856-37
C 15	14.8	74.0	1299	4	US-09-489-039A-738
C 16	14.8	74.0	1982	4	US-08-221-017B-1068
C 17	14.8	74.0	2850	2	US-08-224-482-7
C 18	14.8	74.0	2922	4	US-09-489-039A-4404
C 19	14.8	74.0	3465	4	US-09-489-039A-553
C 20	14.8	74.0	3939	4	US-09-489-039A-4373
C 21	14.4	72.0	2907	4	US-09-620-312D-511
C 22	14.4	72.0	31208	4	US-09-852-067-3
C 23	14.2	71.0	58	3	US-09-140-466-10
C 24	14.2	71.0	463	4	US-09-288-116-73
C 25	14.2	71.0	499	4	US-09-889-9143-7
C 26	14.2	71.0	789	4	US-09-489-039A-4398
C 27	14.2	71.0	850	2	US-08-560-398-7

4	US-09-489-039A-4964	873	71.0	14.2	C 28
4	US-09-443-041A-23	911	71.0	14.2	C 29
4	US-09-543-681A-699	948	71.0	14.2	C 30
4	US-09-543-681A-477	1017	71.0	14.2	C 31
4	US-09-328-352-3405	1041	71.0	14.2	C 32
4	US-09-328-352-1296	1152	71.0	14.2	C 33
4	US-09-489-039A-5098	1533	71.0	14.2	C 34
3	US-08-591-685-8	1633	71.0	14.2	C 35
3	US-09-008-303-1	2057	71.0	14.2	C 36
4	US-09-489-039A-4929	2175	71.0	14.2	C 37
3	US-08-823-110-5	2724	71.0	14.2	C 38
3	US-08-604-298-5	2724	71.0	14.2	C 39
3	US-09-214-916A-3	2935	71.0	14.2	C 40
3	US-08-855-146-1	3321	71.0	14.2	C 41
3	US-08-855-146-3	3321	71.0	14.2	C 42
3	US-08-823-110-2	3833	71.0	14.2	C 43
3	US-08-604-298-2	3833	71.0	14.2	C 44
4	US-09-889-914B-1	3941	71.0	14.2	C 45
4	US-08-961-527-253	4010	71.0	14.2	C 46
4	US-08-961-527-293	4398	71.0	14.2	C 47
4	US-09-214-916A-1	4446	71.0	14.2	C 48
4	US-09-976-594-308	5252	71.0	14.2	C 49
4	US-09-976-594-820	5712	71.0	14.2	C 50
4	US-09-919-172-97	10432	71.0	14.2	C 51
4	US-09-976-594-21	10432	71.0	14.2	C 52
4	US-09-816-093-3	34185	71.0	14.2	C 53
4	US-09-545-481-3	46718	71.0	14.2	C 54
3	US-08-947-823-1	51952	71.0	14.2	C 55
4	US-08-918-421B-1	1664976	71.0	14.2	C 56
4	US-09-557-884-1	1830121	71.0	14.2	C 57
4	US-09-643-990A-1	1830121	71.0	14.2	C 58
3	US-08-687-590-54	1812	70.0	14	C 59
4	US-09-620-312D-65	2155	70.0	14	C 60
4	US-09-566-921-62	2421	70.0	14	C 61
3	US-09-039-859-8	3420	70.0	14	C 62
4	US-09-711-164-106	142	69.0	13.8	C 63
4	US-09-107-532A-1603	330	69.0	13.8	C 64
4	US-09-536-059-21	399	69.0	13.8	C 65
4	US-09-247-155-137	419	69.0	13.8	C 66
4	US-09-247-155-73	425	69.0	13.8	C 67
4	US-09-621-976-1717	440	69.0	13.8	C 68
4	US-09-621-976-19139	466	69.0	13.8	C 69
4	US-09-370-838-126	509	69.0	13.8	C 70
4	US-09-621-976-17716	536	69.0	13.8	C 72
4	US-09-621-976-1822	536	69.0	13.8	C 73
3	US-09-385-982-313	674	69.0	13.8	C 74
4	US-09-489-039A-1016	681	69.0	13.8	C 75
4	US-09-702-705-192	687	69.0	13.8	C 76
4	US-09-736-457-192	687	69.0	13.8	C 77
4	US-09-614-124B-192	687	69.0	13.8	C 78
4	US-09-671-325-192	687	69.0	13.8	C 79
4	US-09-589-184-192	687	69.0	13.8	C 80
4	US-09-853-768-12	782	69.0	13.8	C 81
4	US-09-671-317-284	1001	69.0	13.8	C 82
4	US-08-847-296B-1	1065	69.0	13.8	C 83
4	US-08-567-882-6	1071	69.0	13.8	C 84
4	US-08-720-565-3	1193	69.0	13.8	C 85
4	US-09-016-434-1085	1201	69.0	13.8	C 86
4	US-09-023-655-905	1201	69.0	13.8	C 87
4	US-09-489-039A-2558	1254	69.0	13.8	C 88
4	US-09-489-039A-2522	1348	69.0	13.8	C 89
4	US-09-489-039A-2522	1404	69.0	13.8	C 90
4	US-09-149-476-99	1416	69.0	13.8	C 91
4	US-09-328-352-2627	1482	69.0	13.8	C 92
3	US-09-360-197-9	1602	69.0	13.8	C 93
4	US-08-720-565-1	1689	69.0	13.8	C 94
4	US-09-023-655-959	1717	69.0	13.8	C 95
3	US-09-082-310-4	1760	69.0	13.8	C 96
4	US-09-575-205-4	1760	69.0	13.8	C 97
4	US-09-620-312D-516	1823	69.0	13.8	C 98
3	US-08-575-967A-1	1912	69.0	13.8	C 99
4	US-09-221-017B-1027	2135	69.0	13.8	C 100



```
/ APPLICATION NUMBER: P2911
/ FILING DATE: 09-APR-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/AJ98/01023
/ FILING DATE: 10-DEC-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MONROY, Gladys H
/ REGISTRATION NUMBER: 32,430
/ REFERENCE/DOCKET NUMBER: 27340-20021.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-813-5600
/ TELEFAX: 650-494-0792
/ TELEX: 706141
/ INFORMATION FOR SEQ ID NO: 995:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4400 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: UNKNOWN
/ ORIGINAL SOURCE:
/ ORGANISM: PORYPHYROMONAS GINGIVALIS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1...4400
/
US-09-221-0178-995

Query Match 79.0%; Score 15.8; DB 4; Length 4400;
Best Local Similarity 89.5%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20
|||||
Db 324 TAAAGAGCGCACCAATCC 306

RESULT 3
US-09-489-039A-792/c
/ Sequence 792, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ PRIOR FILING DATE: 2000-01-27
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 792
/ LENGTH: 1476
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-792

Query Match 76.0%; Score 15.2; DB 4; Length 1476;
Best Local Similarity 85.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
|||||
Db 1319 TTAAGAGCGCTCCAAACC 1300

RESULT 4
US-09-976-594-291
/ Sequence 291, Application US/09976594
/ Patent No. 6673549
/ GENERAL INFORMATION:
/ APPLICANT: Furness, Michael
```

```
/ APPLICANT: Buchbinder, Jenny
/ TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
/ FILE REFERENCE: PA-0041 US
/ CURRENT APPLICATION NUMBER: US/09/976,594
/ PRIOR FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: 60/240,409
/ PRIOR FILING DATE: 2000-10-12
/ NUMBER OF SEQ ID NOS: 1143
/ SOFTWARE: PERL Program
/ SEQ ID NO 291
/ LENGTH: 3184
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6673549 333542.1
US-09-976-594-291

Query Match 76.0%; Score 15.2; DB 4; Length 3184;
Best Local Similarity 85.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
|||||
Db 2068 TTAACACCTCTCCAAAGCC 2087

RESULT 5
US-08-849-480A-1
/ Sequence 1, Application US/08849480A
/ Patent No. 5981184
/ GENERAL INFORMATION:
/ APPLICANT: MELCHERS, Klaus
/ TITLE OF INVENTION: SCREENING MODEL
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
/ STREET: 400 - 7th Street, N. W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/849,480A
/ FILING DATE: 02-JUN-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/EP95/04711
/ FILING DATE: 30-NOV-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE P4442970.3
/ FILING DATE: 02-DEC-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE 19505645.0
/ FILING DATE: 18-FEB-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: AISENBERG, Irwin M.
/ REGISTRATION NUMBER: 19,007
/ REFERENCE/DOCKET NUMBER: 8125/P6098AUS0
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202/638-6666
/ TELEFAX: 202/393-5350
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3411 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
```

```
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Helicobacter pylori
/ STRAIN: Helicobacter pylori 69A
/ INDIVIDUAL ISOLATE: Clinical isolate 69A
/ IMMEDIATE SOURCE:
/ LIBRARY: Helicobacter pylori 69A - gene library in
/ LIBRARY: vector pRH160
/ CLONE: pRH439
US-08-849-480A-1

Query Match          76.0%; Score 15.2; DB 2; Length 3411;
Best Local Similarity 85.0%; Pred. No. 70;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| |||||
Db 2927 TTAAGAGCGCTATAAAGCC 2946

RESULT 6
US-09-167-681-29/c
; Sequence 29, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, M.D., Richard M.
; APPLICANT: Raftogiannis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Ottewill, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 7152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3810) ... (3956)
; NAME/KEY: CDS
; LOCATION: (4061) ... (4186)
; NAME/KEY: CDS
; LOCATION: (4276) ... (4374)
; NAME/KEY: CDS
; LOCATION: (5584) ... (5709)
; NAME/KEY: CDS
; LOCATION: (5805) ... (5900)
; NAME/KEY: CDS
; LOCATION: (6426) ... (6605)
; NAME/KEY: CDS
; LOCATION: (6728) ... (6837)
US-09-167-681-29

Query Match          76.0%; Score 15.2; DB 3; Length 7152;
Best Local Similarity 85.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| |||||
Db 2505 TTAAGAGCTCTCCAAAGCC 2486

RESULT 7
US-09-328-174A-1/c
; Sequence 1, Application US/09328174A
; Patent No. 6448003
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco

/ APPLICANT: Kurth, Janice
/ TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
/ FILE REFERENCE: 4389-6 (formerly SEQ-16P)
/ CURRENT APPLICATION NUMBER: US/09/328,174A
/ CURRENT FILING DATE: 1999-06-08
/ PRIOR APPLICATION NUMBER: 09/328,174
/ PRIOR FILING DATE: 1999-06-08
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 8396
/ TYPE: DNA
/ ORGANISM: H. sapiens
US-09-328-174A-1

Query Match          76.0%; Score 15.2; DB 4; Length 8396;
Best Local Similarity 85.0%; Pred. No. 82;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| |||||
Db 2436 TTAAGTGATCTCCAAAGCC 2417

RESULT 8
US-09-167-681-37/c
; Sequence 37, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, M.D., Richard M.
; APPLICANT: Raftogiannis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Ottewill, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 8397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3730) ... (3879)
; NAME/KEY: CDS
; LOCATION: (3987) ... (4112)
; NAME/KEY: CDS
; LOCATION: (4198) ... (4293)
; NAME/KEY: CDS
; LOCATION: (6088) ... (6213)
; NAME/KEY: CDS
; LOCATION: (6309) ... (6404)
; NAME/KEY: CDS
; LOCATION: (7214) ... (7393)
; NAME/KEY: CDS
; LOCATION: (7516) ... (7629)
US-09-167-681-37

Query Match          76.0%; Score 15.2; DB 3; Length 8409;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| |||||
Db 2437 TTAAGTGATCTCCAAAGCC 2418

RESULT 9
US-08-770-379-18
; Sequence 18, Application US/08770379
```

```
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-770-379-18

Query Match 76.0%; Score 15.2; DB 2; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
DB 23242 TTAAGCGCGCTTCAAAGCC 23261

RESULT 10
US-08-757-669A-18
; Sequence 18, Application US/08/757669A
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-757-669A-18

Query Match 76.0%; Score 15.2; DB 3; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
DB 23242 TTAAGCGCGCTTCAAAGCC 23261

RESULT 11
US-09-230-371A-18
; Sequence 18, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 35100
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
; US-09-230-371A-18

Query Match 76.0%; Score 15.2; DB 4; Length 35100;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
DB 23242 TTAAGCGCGCTTCAAAGCC 23261

RESULT 12
US-09-523-656-36
; Sequence 36, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VII-
```

; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707  
; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide  
US-09-523-656-36

Query Match 75.0%; Score 15; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAGCGCTCCAAAGCC 20  
|||||  
DB 9 GAGCGCTCCAAAGCC 23

RESULT 13  
US-09-523-656-33  
; Sequence 33, Application US/09523656  
; Patent No. 6458563  
; GENERAL INFORMATION:  
; APPLICANT: Lollar S., John  
; TITLE OF INVENTION: MODIFIED FACTOR VIII  
; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; CURRENT FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707  
; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 33  
; LENGTH: 105  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:linker  
US-09-523-656-33

Query Match 75.0%; Score 15; DB 4; Length 105;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAGCGCTCCAAAGCC 20  
|||||  
DB 48 GAGCGCTCCAAAGCC 62

RESULT 14  
US-09-523-656-37  
; Sequence 37, Application US/09523656  
; Patent No. 6458563  
; GENERAL INFORMATION:  
; APPLICANT: Lollar S., John  
; TITLE OF INVENTION: MODIFIED FACTOR VIII  
; FILE REFERENCE: 75-951  
; CURRENT APPLICATION NUMBER: US/09/523,656  
; CURRENT FILING DATE: 2000-03-10  
; EARLIER APPLICATION NUMBER: 09/037,601  
; EARLIER FILING DATE: 1998-03-10  
; EARLIER APPLICATION NUMBER: 08/670,707

; EARLIER FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 37  
; LENGTH: 4404  
; TYPE: DNA  
; ORGANISM: Pc-cine  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(4401)  
US-09-523-656-37

Query Match 75.0%; Score 15; DB 4; Length 4404;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAGCGCTCCAAAGCC 20  
|||||  
DB 2310 GAGCGCTCCAAAGCC 2324

RESULT 15  
US-09-489-039A-733/c  
; Sequence 738, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 738  
; LENGTH: 1299  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-738

Query Match 74.0%; Score 14.8; DB 4; Length 1299;  
Best Local Similarity 88.9%; Pred. No. 97;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19  
|||||  
DB 859 TAAAGAGCGCTCCAAAGC 842

RESULT 16  
US-09-221-017B-1068  
; Sequence 1068, Application US/09221017B  
; Patent No. 6444793  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P21182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P21546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P22911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1068:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1...1982
; US-09-221-017B-1068

Query Match 74.0%; Score 14.8; DB 4; Length 1982;
Best Local Similarity 88.9%; Pred. No. 1.e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGC 20
DB 1649 AAAGAGCACTACAAAGC 1666

RESULT 17
US-08-224-482-7
; Sequence 7, Application US/08224482
; Patent No. 5837692
; GENERAL INFORMATION:
; APPLICANT: Mercola, Dan
; TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
; TITLE OF INVENTION: PDGF by Mammalian EGR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/224,482
; FILING DATE: 07-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-ME 9913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2850 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 358..1519
; US-08-224-482-7

Query Match 74.0%; Score 14.8; DB 2; Length 2850;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19
DB 1949 TAAAGAGCGCAACAAAGC 1966

RESULT 18
US-09-489-039A-4404
; Sequence 4404, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4404
; LENGTH: 2922
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-4404

Query Match 74.0%; Score 14.8; DB 4; Length 2922;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19
DB 54 TAAAGCGCGCTTCAAAGC 71

RESULT 19
US-09-489-039A-553
; Sequence 553, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 553
; LENGTH: 3465
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
```



US-09-489-039A-553

Query Match 74.0%; Score 14.8; DB 4; Length 3455;  
Best Local Similarity 82.9%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19

Db 2232 TAAAGAGCGCATCAAGC 2249

RESULT 20

US-09-489-039A-4373/c  
; Sequence 4373, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4373  
; LENGTH: 3939  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4373

Query Match 74.0%; Score 14.8; DB 4; Length 3939;  
Best Local Similarity 83.9%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGC 19

Db 3559 TAAAGCGCGTTCAAAGC 3542

RESULT 21

US-09-620-312D-511/c  
; Sequence 511, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 03/552,317  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: 03/468,725  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: P-FL\_genes Version 1.0

; SEQ ID NO 511  
; LENGTH: 2907  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (359)..(2530)  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2907)  
; OTHER INFORMATION: n = a,t,c or g  
US-09-620-312D-511

Query Match 72.0%; Score 14.4; DB 4; Length 2907;  
Best Local Similarity 93.8%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AGAGCGCTCCAAAGC 20

Db 796 AGAGCGCACCAAGC 781

RESULT 22

US-09-852-067-3  
; Sequence 3, Application US/09852067  
; Patent No. 6531297  
; GENERAL INFORMATION:  
; APPLICANT: MERKULOV, Gennady et al  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
; FILE REFERENCE: CL000897-CIP  
; CURRENT APPLICATION NUMBER: US/09/852,067  
; CURRENT FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 31208  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(31208)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-852-067-3

Query Match 72.0%; Score 14.4; DB 4; Length 31208;  
Best Local Similarity 93.8%; Pred. No. 2.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAA 17

Db 18168 TAAAGAGCGTCCAAA 18183

RESULT 23

US-09-140-466-10  
; Sequence 10, Application US/09140466  
; Patent No. 6268160  
; GENERAL INFORMATION:  
; APPLICANT: CLOUGH, BARBARA  
; APPLICANT: FREISER, PETER  
; APPLICANT: WILSON, ROBERT  
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE  
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS  
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS  
; FILE REFERENCE: N68837S GCW PJC DP  
; CURRENT APPLICATION NUMBER: US/09/140,466  
; CURRENT FILING DATE: 1998-08-26  
; EARLIER APPLICATION NUMBER: US 60/056,246  
; EARLIER FILING DATE: 1997-08-28  
; NUMBER OF SEQ ID NOS: 14

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:48:43 ; Search time 122.706 Seconds

(without alignments)  
729.318 Million cell updates/sec

Title: US-09-877-819B-36

Perfect score: 20

Sequence: 1 ttaagagcgtccaaagcc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Yatch 0%

Maximum Yatch 100%

Lasting first 180 summaries

Database :

Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	20	10	US-09-877-819B-3
C 2	20	100.0	20	10	US-09-877-819B-4
C 3	20	100.0	20	10	US-09-877-819B-36
C 4	20	100.0	20	10	US-09-877-819B-38
C 5	16.8	84.0	494	13	US-10-027-632-352
C 6	16.8	84.0	494	16	US-10-027-632-352
C 7	16.8	84.0	1018	13	US-10-425-114-1011
C 8	16.8	84.0	1299	13	US-10-425-114-20063
C 9	16.8	84.0	2382	13	US-10-424-599-132396
C 10	16.4	82.0	434	13	US-10-424-599-79300
C 11	16	80.0	1800	8	US-08-973-028-3
C 12	15.8	79.0	420	13	US-10-424-599-6560
C 13	15.8	79.0	820	13	US-10-027-632-173218
C 14	15.8	79.0	820	16	US-10-027-632-173218

C 15	15.8	79.0	1110	9	US-09-815-242-7264	Sequence 7264, Ap
C 16	15.8	79.0	1110	13	US-10-282-122A-22599	Sequence 22599, A
C 17	15.8	79.0	1170	10	US-09-882-227-89	Sequence 89, Appl
C 18	15.8	79.0	1581	13	US-10-282-122A-39670	Sequence 39670, A
C 19	15.8	79.0	4400	13	US-10-194-163-995	Sequence 995, App
C 20	15.8	79.0	8546	9	US-09-070-927A-146	Sequence 146, App
C 21	15.8	79.0	68571	16	US-10-401-194-1	Sequence 1, Appl
C 22	15.8	79.0	3309400	9	US-09-738-626-1	Sequence 3971, Ap
C 23	15.4	77.0	732	13	US-10-335-977-3971	Sequence 3971, Ap
C 24	15.4	77.0	735	13	US-10-335-977-3970	Sequence 17, Appl
C 25	15.4	77.0	1176	14	US-10-001-873-17	Sequence 149, App
C 26	15.4	77.0	1440	9	US-09-895-913A-149	Sequence 792, App
C 27	15.4	77.0	10766	10	US-09-764-872-792	Sequence 2971, Ap
C 28	15.2	76.0	187	9	US-09-867-701-2371	Sequence 41104, A
C 29	15.2	76.0	236	13	US-10-424-599-41104	Sequence 1558, Ap
C 30	15.2	76.0	253	13	US-10-085-783A-1558	Sequence 1558, Ap
C 31	15.2	76.0	253	16	US-10-242-535A-1558	Sequence 327, App
C 32	15.2	76.0	265	9	US-09-923-876-327	Sequence 327, App
C 33	15.2	76.0	265	11	US-09-923-876-327	Sequence 327, App
C 34	15.2	76.0	270	13	US-10-335-977-1774	Sequence 1774, Ap
C 35	15.2	76.0	273	13	US-10-335-977-1773	Sequence 1773, Ap
C 36	15.2	76.0	366	13	US-10-424-599-513	Sequence 513, App
C 37	15.2	76.0	414	10	US-09-918-985-5244	Sequence 5244, Ap
C 38	15.2	76.0	456	13	US-10-424-599-72018	Sequence 72018, A
C 39	15.2	76.0	526	13	US-10-424-599-141450	Sequence 141450, A
C 40	15.2	76.0	555	13	US-10-027-632-259306	Sequence 259306, A
C 41	15.2	76.0	555	13	US-10-027-632-259307	Sequence 259307, A
C 42	15.2	76.0	555	16	US-10-027-632-259306	Sequence 259306, A
C 43	15.2	76.0	555	16	US-10-027-632-259307	Sequence 259307, A
C 44	15.2	76.0	570	15	US-10-029-386-9951	Sequence 9951, Ap
C 45	15.2	76.0	615	13	US-10-335-977-2171	Sequence 2171, Ap
C 46	15.2	76.0	747	9	US-09-738-626-3026	Sequence 3026, Ap
C 47	15.2	76.0	762	9	US-09-974-300-1956	Sequence 1956, Ap
C 48	15.2	76.0	765	13	US-10-335-977-416	Sequence 416, App
C 49	15.2	76.0	768	13	US-10-335-977-415	Sequence 415, App
C 50	15.2	76.0	787	13	US-10-424-599-105581	Sequence 105581, A
C 51	15.2	76.0	807	13	US-10-335-977-1776	Sequence 1776, Ap
C 52	15.2	76.0	948	16	US-10-369-493-46752	Sequence 46752, A
C 53	15.2	76.0	1110	9	US-09-815-242-7432	Sequence 7432, Ap
C 54	15.2	76.0	1287	13	US-10-335-977-782	Sequence 782, App
C 55	15.2	76.0	176	13	US-10-424-599-105582	Sequence 105582, A
C 56	15.2	76.0	1840	13	US-10-424-599-89421	Sequence 89421, A
C 57	15.2	76.0	1876	16	US-10-369-493-33650	Sequence 33650, A
C 58	15.2	76.0	2052	16	US-10-369-493-35696	Sequence 35696, A
C 59	15.2	76.0	2061	9	US-09-815-242-7440	Sequence 7440, Ap
C 60	15.2	76.0	2253	9	US-09-738-626-3025	Sequence 3025, Ap
C 61	15.2	76.0	2385	13	US-10-027-632-112010	Sequence 112010, A
C 62	15.2	76.0	2385	16	US-10-027-632-112010	Sequence 112010, A
C 63	15.2	76.0	2550	16	US-10-108-260A-2360	Sequence 2360, Ap
C 64	15.2	76.0	3420	13	US-09-795-651-49	Sequence 49, Appl
C 65	15.2	76.0	5945	13	US-10-424-599-105261	Sequence 105261, A
C 66	15.2	76.0	7152	13	US-09-829-545-29	Sequence 29, Appl
C 67	15.2	76.0	8396	15	US-10-206-839-1	Sequence 1, Appl
C 68	15.2	76.0	8397	13	US-09-829-545-37	Sequence 37, Appl
C 69	15.2	76.0	193853	13	US-10-087-152-1663	Sequence 1663, Ap
C 70	15.2	76.0	251364	15	US-10-175-523-58	Sequence 61, Appl
C 71	15.2	76.0	251364	15	US-10-175-523-79	Sequence 79, Appl
C 72	15.2	76.0	251364	15	US-10-027-632-123862	Sequence 123862, A
C 73	15	75.0	1238	13	US-10-027-632-123862	Sequence 123862, A
C 74	15	75.0	1238	16	US-10-027-632-123862	Sequence 9827, A
C 75	14.8	74.0	267	13	US-10-424-599-98627	Sequence 8240, Ap
C 76	14.8	74.0	301	9	US-09-974-300-8240	Sequence 10521, A
C 77	14.8	74.0	352	9	US-09-864-761-10521	Sequence 153, App
C 78	14.8	74.0	360	16	US-10-259-194A-153	Sequence 24462, A
C 79	14.8	74.0	386	9	US-09-864-761-24462	Sequence 27143, A
C 80	14.8	74.0	408	9	US-09-864-761-27143	Sequence 51, Appl
C 81	14.8	74.0	480	9	US-09-772-134B-51	Sequence 52, Appl
C 82	14.8	74.0	480	9	US-09-772-134B-52	Sequence 7758, Ap
C 83	14.8	74.0	534	9	US-09-864-761-7758	Sequence 125839, A
C 84	14.8	74.0	583	13	US-10-027-632-125839	Sequence 125840, A
C 85	14.8	74.0	583	13	US-10-027-632-125840	Sequence 125841, A
C 86	14.8	74.0	583	13	US-10-027-632-125841	Sequence 125839, A
C 87	14.8	74.0	583	16	US-10-027-632-125839	

```
C 88 14.8 74.0 583 16 US-10-027-632-125840 Sequence 125840,
C 89 14.8 74.0 583 16 US-10-027-632-125841 Sequence 125841,
C 90 14.8 74.0 605 13 US-10-027-632-89299 Sequence 89299, A
14.8 74.0 605 13 US-10-027-632-89300 Sequence 89300, A
14.8 74.0 605 13 US-10-027-632-317126 Sequence 317126,
14.8 74.0 605 13 US-10-027-632-317127 Sequence 317127,
14.8 74.0 605 13 US-10-027-632-317128 Sequence 317128,
14.8 74.0 605 16 US-10-027-632-89299 Sequence 89299, A
14.8 74.0 605 16 US-10-027-632-89300 Sequence 89300, A
14.8 74.0 605 16 US-10-027-632-89301 Sequence 89301, A
14.8 74.0 605 16 US-10-027-632-317126 Sequence 317126,
14.8 74.0 605 16 US-10-027-632-317127 Sequence 317127,
14.8 74.0 605 16 US-10-027-632-317128 Sequence 317128,
14.8 74.0 745 13 US-10-424-599-70431 Sequence 70431, A
14.8 74.0 765 9 US-09-738-626-75 Sequence 75, Appl
14.8 74.0 777 16 US-10-398-221-741 Sequence 741, Appl
14.8 74.0 818 13 US-10-424-599-63507 Sequence 63507, A
14.8 74.0 912 13 US-10-027-632-9130 Sequence 9130, Ap
14.8 74.0 972 16 US-10-027-632-9130 Sequence 9130, Ap
14.8 74.0 1050 15 US-10-239-431A-5 Sequence 5, Appli
14.8 74.0 1128 15 US-10-239-431A-3 Sequence 3, Appli
14.8 74.0 1156 9 US-09-917-800A-1347 Sequence 1347, Ap
14.8 74.0 1156 15 US-10-205-194-32 Sequence 32, Appl
14.8 74.0 1156 16 US-10-388-934-159 Sequence 159, Ap
14.8 74.0 1467 9 US-09-738-626-76 Sequence 76, Appl
14.8 74.0 1830 15 US-10-239-431A-1 Sequence 1, Appli
14.8 74.0 1920 10 US-09-989-643-27 Sequence 27, Appl
14.8 74.0 1920 16 US-10-398-221-743 Sequence 743, Ap
14.8 74.0 1920 16 US-10-398-221-2645 Sequence 2645, Ap
14.8 74.0 1982 13 US-10-194-163-1068 Sequence 1068, Ap
14.8 74.0 2610 16 US-10-369-493-37578 Sequence 37578, A
14.8 74.0 2631 13 US-10-282-122A-41142 Sequence 41142, A
14.8 74.0 2726 9 US-09-764-864-7722 Sequence 772, Ap
14.8 74.0 2744 13 US-09-764-864-357 Sequence 357, Ap
14.8 74.0 3074 9 US-10-282-122A-19458 Sequence 19458, A
14.8 74.0 3444 13 US-10-282-122A-23565 Sequence 23565, A
14.8 74.0 3885 13 US-10-282-122A-23663 Sequence 23663, A
14.8 74.0 4272 10 US-09-960-706-1074 Sequence 1074, Ap
14.8 74.0 4338 15 US-10-247-671-104 Sequence 104, Ap
14.8 74.0 4339 14 US-10-002-600-50 Sequence 50, Appl
14.8 74.0 145831 9 US-09-969-708-79 Sequence 79, Appl
14.8 74.0 145831 9 US-09-954-456-2116 Sequence 2116, Ap
14.8 74.0 145831 10 US-09-873-3670-646 Sequence 646, Ap
14.8 74.0 145831 13 US-10-240-425-363 Sequence 363, Ap
14.8 74.0 653122 13 US-10-087-192-225 Sequence 226, Ap
14.8 74.0 3309400 9 US-09-738-626-1 Sequence 1, Appli
14.8 74.0 1296 9 US-09-864-761-21627 Sequence 21627, A
14.8 74.0 186 13 US-10-282-122A-11502 Sequence 11502, A
14.8 74.0 195 13 US-10-282-122A-1244 Sequence 1244, A
14.8 74.0 213 9 US-09-864-761-21445 Sequence 21445, A
14.8 74.0 268 13 US-10-424-599-72721 Sequence 72721, A
14.8 74.0 309 13 US-10-085-783A-56225 Sequence 56225, A
14.8 74.0 309 16 US-10-242-535A-56225 Sequence 4705, Ap
14.8 74.0 426 9 US-09-864-761-4705 Sequence 4893, Ap
14.8 74.0 435 9 US-09-864-761-4893 Sequence 4893, Ap
14.8 74.0 474 10 US-09-318-995-29217 Sequence 29217, A
14.8 74.0 504 13 US-10-243-552-224 Sequence 224, Ap
14.8 74.0 511 13 US-10-424-599-93201 Sequence 93201, A
14.8 74.0 536 16 US-10-341-961A-2320 Sequence 253, Ap
14.8 74.0 581 13 US-10-027-632-15230 Sequence 15230, A
14.8 74.0 581 16 US-10-027-632-15230 Sequence 15230, A
14.8 74.0 609 13 US-10-335-977-3577 Sequence 3577, Ap
14.8 74.0 611 13 US-10-027-632-229898 Sequence 229898,
14.8 74.0 611 16 US-10-027-632-229898 Sequence 409, Ap
14.8 74.0 663 9 US-09-841-132-409 Sequence 1149, Ap
14.8 74.0 1251 13 US-10-335-977-1149 Sequence 1150, Ap
14.8 74.0 1265 13 US-10-335-977-1150 Sequence 1151, Ap
14.8 74.0 1278 13 US-10-335-977-1151 Sequence 109, Ap
14.8 74.0 1508 13 US-10-302-172-109 Sequence 41372, A
14.8 74.0 2119 13 US-10-282-122A-41372 Sequence 2502, Ap
14.8 74.0 2258 13 US-10-335-977-2502 Sequence 34, Appl
14.8 74.0 2456 13 US-09-939-853A-34 Sequence 36, Appl
14.8 74.0 2456 13 US-09-939-853A-36 Sequence 2503, Ap
14.8 74.0 2547 13 US-10-335-977-2503 Sequence 2504, Ap
14.8 74.0 2547 13 US-10-335-977-2504 Sequence 635, Ap
14.8 74.0 2547 13 US-10-104-047-635 Sequence 24513, A
14.8 74.0 2547 13 US-10-282-122A-24913 Sequence 9475, Ap
14.8 74.0 2547 13 US-09-764-891-9475 Sequence 9476, Ap
14.8 74.0 2547 13 US-09-764-891-9475 Sequence 856, Ap
14.8 74.0 2547 13 US-10-205-428-856 Sequence 511, Ap
14.8 74.0 2547 13 US-10-037-270-511 Sequence 511, Ap
14.8 74.0 2547 13 US-10-117-722-511 Sequence 511, Ap
14.8 74.0 2547 13 US-10-017-161-709 Sequence 709, Ap
14.8 74.0 2547 13 US-10-292-798-621 Sequence 621, Ap
14.8 74.0 2547 13 US-09-852-067-3 Sequence 3, Appli
14.8 74.0 2547 13 US-10-338-691-3 Sequence 3, Appli
14.8 74.0 2547 13 US-10-087-192-1147 Sequence 1147, Ap
14.8 74.0 2547 13 US-10-002-491-10 Sequence 10, Appli
14.8 74.0 2547 13 US-10-067-514-1 Sequence 1, Appli
14.8 74.0 2547 13 US-10-419-723-1 Sequence 1, Appli
14.8 74.0 2547 13 US-09-845-335-10 Sequence 10, Appli
14.8 74.0 2547 13 US-10-025-386-22085 Sequence 22085, A
14.8 74.0 2547 13 US-09-864-761-21439 Sequence 21439, A
```

## ALIGNMENTS

## RESULT 1

```
US-09-877-819B-3/c
; Sequence 3, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Address tag
US-09-877-819B-3
Query Match 100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 TTAAGAGCGCTCCAAAGCC 20

Db 20 TTAAGAGCGCTCCAAAGCC 1

## RESULT 2

```
US-09-877-819B-4
; Sequence 4, Application US/09877819B
; Publication No. US20030190609A1
; GENERAL INFORMATION:
; APPLICANT: Torney, David
; TITLE OF INVENTION: Address/Capture Tags For Flow-Cytometry Based Minisequencing
; FILE REFERENCE: S-94,664
; CURRENT APPLICATION NUMBER: US/09/877,819B
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
```

OTHER INFORMATION: Capture tag  
US-09-877-819B-4

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 TTAAGAGCGCTCCAAAGCC 20

## RESULT 3

US-09-877-819B-36  
Sequence 36, Application US/09877819B  
Publication No. US20030190609A1  
GENERAL INFORMATION:  
APPLICANT: White, Scott  
TITLE OF INVENTION: Torney, David  
FILE REFERENCE: S-94,664  
CURRENT APPLICATION NUMBER: US/09/877,819B  
CURRENT FILING DATE: 2001-06-07  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 36  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Capture sequence  
US-09-877-819B-36

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 TTAAGAGCGCTCCAAAGCC 20

## RESULT 4

US-09-877-819B-38/c  
Sequence 38, Application US/09877819B  
Publication No. US20030190609A1  
GENERAL INFORMATION:  
APPLICANT: White, Scott  
TITLE OF INVENTION: Torney, David  
FILE REFERENCE: S-94,664  
CURRENT APPLICATION NUMBER: US/09/877,819B  
CURRENT FILING DATE: 2001-06-07  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 38  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Address sequence  
US-09-877-819B-38

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 20 TTAAGAGCGCTCCAAAGCC 1

## RESULT 5

US-10-027-632-3532/c  
Sequence 3532, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3532  
LENGTH: 494  
TYPE: DNA  
ORGANISM: Human  
OTHER INFORMATION: Capture sequence  
US-10-027-632-3532

Query Match 84.0%; Score 16.8; DB 13; Length 494;  
Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGCC 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 334 TCAAGAGTCTCCAAAGCC 315

## RESULT 6

US-10-027-632-3532/c  
Sequence 3532, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3532  
LENGTH: 494  
TYPE: DNA  
ORGANISM: Human  
OTHER INFORMATION: Capture sequence  
US-10-027-632-3532

```
Query Match      84.0%; Score 16.8; DB 16; Length 494;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| ||||| |||||
Db 334 TCAAGAGTGCTCCAAAGCC 315

RESULT 7
US-10-425-114-10111/c
; Sequence 10111, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10111
; LENGTH: 1018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700894185_FLI
US-10-425-114-10111

Query Match      84.0%; Score 16.8; DB 13; Length 1018;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| ||||| |||||
Db 854 TTAAGAGGCTTCCAAAGCC 835

RESULT 8
US-10-425-114-20063/c
; Sequence 20063, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20063
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3170-053-F12_FLI
US-10-425-114-20063

Query Match      84.0%; Score 16.8; DB 13; Length 1299;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| ||||| |||||
Db 517 TTAAGAGAGGCTTCAAGCC 498

RESULT 9
US-10-424-599-132396/c
; Sequence 132396, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132396
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90563C.1
US-10-424-599-132396

Query Match      84.0%; Score 16.8; DB 13; Length 2382;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20
    ||||| ||||| ||||| |||||
Db 2150 TTAAGAGGCTTCCAAAGCC 2131

RESULT 10
US-10-424-599-79300
; Sequence 79300, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 79300
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42623C.1
US-10-424-599-79300

Query Match      82.0%; Score 16.4; DB 13; Length 434;
Best Local Similarity 94.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTAAGAGCGCTCCAAAGCC 19
    ||||| ||||| ||||| |||||
Db 340 TAAAGAGGCTCCAAAGC 357

RESULT 11
US-38-973-028-3
; Sequence 3, Application US/08973028
; Publication No. US20020028210A1
```

GENERAL INFORMATION:  
APPLICANT: Berglindh, Thomas  
APPLICANT: Bolin, Ingrid  
APPLICANT: Mellgard, Bjorn  
APPLICANT: Svennerholm, Ann-Mari  
TITLE OF INVENTION: A Vaccine Composition Comprising Helicobacter  
TITLE OF INVENTION: Pylori Flagellin Polypeptide  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case LLP  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (220)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,028  
FILING DATE: 03-DEC-1997  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE97/01928  
FILING DATE: 18-NOV-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: SE 9604322-9  
FILING DATE: 25-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cleland, Thelma A. Chen  
REGISTRATION NUMBER: 40,948  
REFERENCE/DOCKET NUMBER: 1103326-0289  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8515  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1800 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 138..1682  
OTHER INFORMATION: /product= "FlaB protein"  
US-08-973-028-3

Query Match 80.0%; Score 16; DB 8; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 AAGAGCGCTCCAAAGC 19  
Db 464 AAGAGCGCTCCAAAGC 479  
RESULT 12  
US-10-424-599-6560/c  
Sequence 6560, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 6560  
LENGTH: 790  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_105933C.1  
US-10-424-599-6560

Query Match 79.0%; Score 15.8; DB 13; Length 790;  
Best Local Similarity 89.5%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20  
Db 621 TAAAGAGCGCTCCAAAGCC 603

RESULT 13  
US-10-027-632-173218/c  
Sequence 173218, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIORITY APPLICATION NUMBER: US 60/218,006  
PRIORITY FILING DATE: 2000-07-12  
PRIORITY APPLICATION NUMBER: US 60/198,676  
PRIORITY FILING DATE: 2000-04-20  
PRIORITY APPLICATION NUMBER: US 60/193,483  
PRIORITY FILING DATE: 2000-03-29  
PRIORITY APPLICATION NUMBER: US 60/185,218  
PRIORITY FILING DATE: 2000-02-24  
PRIORITY APPLICATION NUMBER: US 60/167,363  
PRIORITY FILING DATE: 1999-11-23  
PRIORITY APPLICATION NUMBER: US 60/156,358  
PRIORITY FILING DATE: 1999-09-28  
PRIORITY APPLICATION NUMBER: US 60/146,002  
PRIORITY FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: PatsEQ for Windows Version 4.0  
SEQ ID NO 173218  
LENGTH: 820  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-173218

Query Match 79.0%; Score 15.8; DB 13; Length 820;  
Best Local Similarity 89.5%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20  
Db 510 TAAAGAGCGCTCCAAAGCC 492

RESULT 14  
US-10-027-632-173218/c  
Sequence 173218, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIORITY APPLICATION NUMBER: US 60/218,006  
PRIORITY FILING DATE: 2000-07-12



```
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. JS20030158396A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24)...(1130)
US-09-882-227-89

Query Match          79.0%; Score 15.8; DB 10; Length 1170;
Best Local Similarity 89.5%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TAAAGAGCGCTCCAAAGCC 20
Db 795 TATAGAGCGCTCCAAAGCC 777

RESULT 18
US-10-282-122A-39670
; Sequence 39670, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39670
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Salmonella typhi
US-10-282-122A-39670

Query Match          79.0%; Score 15.8; DB 13; Length 1581;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TTAAGAGCGCTCCAAAGC 19
Db 1280 TTAAGAGCGCGCAAGC 1298

RESULT 19
US-10-194-163-995/c
; Sequence 995, Application US/10194163
; Publication No. US20020172976A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce Carter
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYNUCLEOTIDES
; AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/194,163
; FILING DATE: 04-Nov-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Basu, Shantanu
; REGISTRATION NUMBER: 43,318
; REFERENCE/DOCKET NUMBER: 5292a2000101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5995
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 995
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...4400
; SEQUENCE DESCRIPTION: SEQ ID NO: 995
US-10-194-163-995

Query Match          79.0%; Score 15.8; DB 13; Length 4400;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TAAAGAGCGCTCCAAAGCC 20
```





ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/335,977  
FILING DATE: 30-Dec-2002  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/993,002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 3971:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...732  
SEQUENCE DESCRIPTION: SEQ ID NO: 3971:  
US-10-335-977-3971

Query Match 77.0%; Score 15.4; DB 13; Length 732;  
Best Local Similarity 94.1%; Pred. No. 6.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AAGAGCGCTCCAAAGCC 20  
|||||  
Db 41 AAGAGCGCACCAAGCC 57

RESULT 24  
US-10-335-977-3970  
Sequence 3970, Application US/10335977  
Publication No. US20040052799A1  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS SMITH et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 10031  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/335,977  
FILING DATE: 30-Dec-2002  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/993,002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 3970:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...735  
SEQUENCE DESCRIPTION: SEQ ID NO: 3970:  
US-10-335-977-3970

Query Match 77.0%; Score 15.4; DB 13; Length 735;  
Best Local Similarity 94.1%; Pred. No. 6.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AAGAGCGCTCCAAAGCC 20  
|||||  
Db 41 AAGAGCGCACCAAGCC 57

RESULT 25  
US-10-001-873-17  
Sequence 17, Application US/10001873  
Publication No. US20020160388A1  
GENERAL INFORMATION:  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Heive  
APPLICANT: Chen, Sei-yu  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
APPLICANT: Turner, Leah  
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
FILE REFERENCE: DEX-0275  
CURRENT APPLICATION NUMBER: US/10/001,873  
CURRENT FILING DATE: 2001-11-20  
PRIORITY APPLICATION NUMBER: 60/252,055  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: 60/252,496  
PRIOR FILING DATE: 2000-11-22  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 1176  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-001-873-17

Query Match 77.0%; Score 15.4; DB 14; Length 1176;  
Best Local Similarity 94.1%; Pred. No. 6.4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AAGAGCGCTCCAAAGCC 20  
|||||  
Db 696 AAGAGCGCACCAAGCC 712

```
RESULT 26
US-09-895-913A-149
; Sequence 149, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Keanthous, Harold
; APPLICANT: Al-Garawi, Anal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)...(1407)
US-09-895-913A-149

Query Match      77.0%; Score 15.4; DB 9; Length 1440;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4  AAGAGCGCTCCAAAGCC 20
Db      443 AAGAGCGCACCAAGCC 459

RESULT 27
US-09-764-872-792/c
; Sequence 792, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL25
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 792
; LENGTH: 10766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-792

Query Match      77.0%; Score 15.4; DB 10; Length 10766;
Best Local Similarity 94.1%; Pred. No. 7.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3  AAAGAGCGCTCCAAAGC 19
Db      9732 AAAGAGCACTCCAAAGC 9716

RESULT 28
US-09-867-701-2971/c
; Sequence 2971, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
```

```
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2971
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-2971

Query Match      76.0%; Score 15.2; DB 9; Length 187;
Best Local Similarity 85.0%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  TTAAGAGCGCTCCAAAGCC 20
Db      60  TTAAGTGCCTTCAAAACC 41

RESULT 29
US-10-424-599-41104/c
; Sequence 41104, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(S3223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 41104
; LENGTH: 236
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(236)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT3847_137117C.1
US-10-424-599-41104

Query Match      76.0%; Score 15.2; DB 13; Length 236;
Best Local Similarity 85.0%; Pred. No. 7.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  TTAAGAGCGCTCCAAAGCC 20
Db      133 TTAGAGAGAGCTCCCAAGCC 114

RESULT 30
US-10-085-783A-1558
; Sequence 1558, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChordroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1272.82 seconds  
(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819B-36

Perfect score: 20

Sequence: 1 ttaaagagcgctccaaagcc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vxl:\*

28: gb\_gssl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	18	90.0	426	9	AI553971	AI553971 t649a10.x
2	18	90.0	516	9	AI910868	AI910868 wd20f03.x
3	18	90.0	545	10	AW188109	AW188109 xj92d12.x
4	18	90.0	891	12	BG717067	BG717067 602689141

597	29	BX223249	Danio rer
615	14	CA354389	626102 NC
631	28	AQ756548	HS_5373_B
685	14	CF835325	UCRC503_0
277	10	BB529596	BB529596
331	10	BE022711	sm87e03.y
366	9	AA569421	Tc-EST-03
423	12	BM092062	sah07b01.
425	12	BG511461	sad02c02.
427	9	AI228821	ES123516
497	14	CF863082	ps25005XN
498	14	CA707288	wk2c.pk0
513	13	BQ299456	sao39h03.
516	10	BE115037	UI-R-BJ1-
536	12	BI1301021	UI-R-DK0-
542	12	BM953746	sm85f03.
577	13	BU546888	GM880010B
586	9	AA603956	Tc-EST-08
609	14	CD891133	G118.116E
636	10	BB480806	BB480806
652	12	BJ624920	BJ624920
654	12	BM107891	BM107891
663	9	AA683466	Tc-EST-04
663	29	CE237493	tigr-gss-
671	12	BG839743	Gm01_02d0
696	13	BY712910	BY712910
704	14	CD898318	G174.108L
723	29	AG185598	Pan tigr1
798	29	CC906028	cc906028
862	14	CE290368	UCRC501_0
870	29	CG954094	MBEPO20TR
932	10	BE906019	601497026
934	10	BF103354	601646785
973	28	AQ743123	HS_5384_B
1004	11	AK013846	Mus muscu
1268	11	AK084835	Mus muscu
1277	11	AK089953	Mus muscu
488	14	CP263777	AUA IPT-K
538	28	AZ164174	SP_0075_A
553	28	AQ725857	HS_2100_B
557	14	CB213545	OML03825
576	12	BJ003949	BJ003949
603	9	AV858696	AV858696
611	9	AV967795	AV967795
651	28	BZ202298	CH230-399
663	13	BM073951	BM073951
693	29	CE632455	tigr-gss-
726	28	AZ330106	1M0055111
839	28	BZ560447	pacs2-164
868	14	CB994896	AGENCOURT
899	29	CNS06VJA	T3 end of
1036	29	CG872551	ZMMEBC027
1142	29	CNS02QHA	Tetraodon
265	10	BF710761	MI-P-AV1-
527	28	CC130113	NDL_76011
570	9	AI730174	BNLGH1587
1002	12	BG429601	602501281
160	29	CE705231	tigr-gss-
247	10	BB568268	BB568268
261	28	AQ001200	CIT-HSP-2
276	28	AZ725377	RPCT-24-1
294	28	AQ383346	RPCT11-13
343	28	CC440381	PJHXY77TD
346	12	BM259540	952010H02
368	10	BE453962	946048E11
376	13	C27262	Rice
386	14	CF336070	JMT--05-P
389	9	AI769023	wg31g07.x
392	13	BY225009	BY225009
397	29	CG633445	OST353099
404	13	BU615908	UI-H-DF0-
412	12	BM500797	BM500797
420	9	AJ475177	AJ475177

C 78	15.8	79.0	441	9	AL663926	AL663926	AL663926	C 151	15.8	79.0	1112	29	AG182916	AG182916	Pan trogl
C 79	15.8	79.0	444	12	N67764	za02f12.81		C 152	15.8	79.0	1124	13	BUI27110	BUI27110	603113933
C 80	15.8	79.0	450	14	BP522369	BP522369		C 153	15.8	79.0	1752	10	BF706623	BF706623	281320 NA
C 81	15.8	79.0	453	9	AU006479	AU006479		C 154	15.4	77.0	148	10	BF706623	BF706623	281320 NA
C 82	15.8	79.0	468	28	AZ402814	IMC170J17		C 155	15.4	77.0	161	29	TA336D08Q	TA336D08Q	T. Bruce1
C 83	15.8	79.0	476	10	AW741895	UG97G09.Y		C 156	15.4	77.0	195	14	CD187527	MS1-0031P	MS1-0031P
C 84	15.8	79.0	480	9	AJ469281	AJ469281		C 157	15.4	77.0	212	14	CD063512	MA1-0031P	MA1-0031P
C 85	15.8	79.0	486	28	BH476864	BOHO175TF		C 158	15.4	77.0	213	10	BF828480	MR1-HN006	MR1-HN006
C 86	15.8	79.0	488	14	CB471135	8S334.12		C 159	15.4	77.0	245	10	BB603751	BB603751	BB603751
C 87	15.8	79.0	491	13	BO741584	89C20501.		C 160	15.4	77.0	267	29	CE752917	ti9t-gss-	ti9t-gss-
C 88	15.8	79.0	492	10	AW654647	104664 NA		C 161	15.4	77.0	288	10	BB080100	BB080100	BB080100
C 89	15.8	79.0	492	29	CC982363	ZUAD106TH		C 162	15.4	77.0	297	14	CD060462	RZ152A1P0	RZ152A1P0
C 90	15.8	79.0	494	12	EM073825	MBS776-DO		C 163	15.4	77.0	328	14	CB923323	VVD095A08	VVD095A08
C 91	15.8	79.0	499	9	AI600575	486062D01		C 164	15.4	77.0	341	14	CD186983	CD186983	CD186983
C 92	15.8	79.0	506	14	D42733	D42733	Rice	C 165	15.4	77.0	343	14	CD187431	CD187431	CD187431
C 93	15.8	79.0	528	28	AQ783813	RS_2001.A		C 166	15.4	77.0	344	14	CD136713	CD136713	CD136713
C 94	15.8	79.0	535	13	BU988978	HF1921F		C 167	15.4	77.0	377	14	CD187393	CD187393	CD187393
C 95	15.8	79.0	536	29	CNS07G33	AL610077	Anopheles	C 168	15.4	77.0	379	14	CD187525	MS1-0056U	MS1-0056U
C 96	15.8	79.0	539	9	AA021775	mh85e08.r		C 169	15.4	77.0	385	14	CD187178	MS1-0056U	MS1-0056U
C 97	15.8	79.0	541	12	BI542689	949021C08		C 170	15.4	77.0	386	14	CD186922	MS1-0056U	MS1-0056U
C 98	15.8	79.0	545	12	BM259212	952010H02		C 171	15.4	77.0	419	14	CD186788	MS1-0056U	MS1-0056U
C 99	15.8	79.0	545	13	BM089196	XZ23808.Y		C 172	15.4	77.0	419	14	CD187441	MS1-0056U	MS1-0056U
C 100	15.8	79.0	548	10	BE475793	SE475793	946048E11	C 173	15.4	77.0	420	14	CD187167	MS1-0056U	MS1-0056U
C 101	15.8	79.0	562	9	AV919296	AV919296		C 174	15.4	77.0	420	14	CD187275	MS1-0056U	MS1-0056U
C 102	15.8	79.0	595	13	BQ832203	LI61n2090		C 175	15.4	77.0	420	14	CD187311	MS1-0056U	MS1-0056U
C 103	15.8	79.0	600	29	AG212944	Oryza sat		C 176	15.4	77.0	420	14	CD187416	MS1-0056U	MS1-0056U
C 104	15.8	79.0	611	9	AV914262	AV914262		C 177	15.4	77.0	420	14	CD187461	MS1-0056U	MS1-0056U
C 105	15.8	79.0	612	28	BZ202953	CH230-304		C 178	15.4	77.0	420	14	CD187465	MS1-0056U	MS1-0056U
C 106	15.8	79.0	619	28	AZ603720	IM0423011		C 179	15.4	77.0	420	14	CD187492	MS1-0056U	MS1-0056U
C 107	15.8	79.0	625	29	CC862698	NDL.64M14		C 180	15.4	77.0	420	14	CD187606	MS1-0056U	MS1-0056U
C 108	15.8	79.0	632	29	CG819239	SOYEL04TH									
C 109	15.8	79.0	639	12	BM485692	pgmlc.pk0									
C 110	15.8	79.0	649	14	CA418535	UI-H-EZ1-									
C 111	15.8	79.0	651	10	BP435588	RF435588	rac33e01.								
C 112	15.8	79.0	655	10	AW505948	GE1215.G1									
C 113	15.8	79.0	658	28	AQ584860	RPC1-11-3									
C 114	15.8	79.0	659	13	CA122050	SCJFLR107									
C 115	15.8	79.0	671	13	CA122050	SCJFLR107									
C 116	15.8	79.0	676	28	CC431341	PUEK42TD									
C 117	15.8	79.0	677	13	BU449519	603765641									
C 118	15.8	79.0	678	9	AI296401	LPI0272.5									
C 119	15.8	79.0	678	14	CA763938	BY762795	BY762795								
C 120	15.8	79.0	686	13	BY762795	BY762795									
C 121	15.8	79.0	688	28	AZ326536	IM0049A03									
C 122	15.8	79.0	690	14	CA189995	SCCRLR1C0									
C 123	15.8	79.0	695	13	BM042141	BM042141									
C 124	15.8	79.0	701	28	AQ006317	AQ006317	CIT-RSP-2								
C 125	15.8	79.0	703	9	AI181066	AI181066	ud72H09.Y								
C 126	15.8	79.0	703	10	BF500388	AT15261.5									
C 127	15.8	79.0	706	13	BU381297	603860066									
C 128	15.8	79.0	712	12	BI358642	GM28559.5									
C 129	15.8	79.0	714	9	AW006171	w293B25.x									
C 130	15.8	79.0	723	28	BH976641	odf79h01.									
C 131	15.8	79.0	724	29	AG169417	AG169417	Pan trogl								
C 132	15.8	79.0	751	13	CA122766	SCULR110									
C 133	15.8	79.0	769	14	CF253965	mdv110.b									
C 134	15.8	79.0	772	13	BU243541	603779751									
C 135	15.8	79.0	772	14	CK291961	BST754675									
C 136	15.8	79.0	800	28	CC096858	CC096858	CSU-K34.1								
C 137	15.8	79.0	803	29	CG951694	MEEAQ16TR									
C 138	15.8	79.0	808	29	CG952126	FUFQC19TD									
C 139	15.8	79.0	809	14	CD115756	ME1-0038P									
C 140	15.8	79.0	817	29	CNS02XNC	AL218589	Tet-raodon								
C 141	15.8	79.0	863	29	CG177055	PUKX34TD									
C 142	15.8	79.0	894	10	BF108273	601924361									
C 143	15.8	79.0	927	13	BU281316	603864148									
C 144	15.8	79.0	939	12	BI112382	602900075									
C 145	15.8	79.0	957	29	CNS01HOM	AL143831	Anopheles								
C 146	15.8	79.0	966	29	CNS02C4C	AL070240	Tetraodon								
C 147	15.8	79.0	972	29	CNS00FTM	CG115157	PufQ280FD								
C 148	15.8	79.0	973	28	CG115157	PufQ280FD									
C 149	15.8	79.0	1038	28	AZ209312	SP_0101.A									
C 150	15.8	79.0	1082	14	CK163088	FGA501570									

## ALIGNMENTS

## RESULT 1

AI553971

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1. 426

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2090010"

/lab\_host="DH10B"

/clone\_lib="Soares NFL T GBC S1"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not 1; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NCT-CCAP GCB1) were mixed, and ss circles were made in

vitro. Following EAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

## ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 426;  
Best Local Similarity 100.0%; Pred. No. 1.9e-02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
|||||  
Db 125 AAAGAGCGCTCCAAAGCC 142

## RESULT 2

AI910868 516 bp mRNA linear EST 17-DEC-1999  
LOCUS wd20f03.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
DEFINITION IMAGE:2328701 3', mRNA sequence.

ACCESSION AI910868  
VERSION AI910868  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 516)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 556, Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 461.  
Location/Qualifiers  
1..516  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2328701"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/notes="Organ: pooled; Vector: pWT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

## FEATURES

source

Query Match 90.0%; Score 18; DB 10; Length 545;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
|||||  
Db 125 AAAGAGCGCTCCAAAGCC 142

## RESULT 4

BG717067/c  
LOCUS 60248941Fl NTH MGC\_97 Homo sapiens cDNA clone IMAGE:4621395 5',  
DEFINITION mRNA sequence.  
ACCESSION BG717067  
VERSION BG717067.1 GI:13996254  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

Query Match 90.0%; Score 18; DB 9; Length 516;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
|||||  
Db 125 AAAGAGCGCTCCAAAGCC 142

## RESULT 3

## AW188109

LOCUS xj92d12.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
DEFINITION IMAGE:2664695 3', mRNA sequence.

ACCESSION AW188109  
VERSION AW188109.1 GI:6462545  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 545)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40UP from Gibco  
High quality sequence stop: 458.  
Location/Qualifiers  
1..545  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2664695"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/notes="Organ: pooled; Vector: pWT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

FEATURES  
source

## ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 545;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
|||||  
Db 125 AAAGAGCGCTCCAAAGCC 142

## RESULT 4

BG717067/c  
LOCUS 60248941Fl NTH MGC\_97 Homo sapiens cDNA clone IMAGE:4621395 5',  
DEFINITION mRNA sequence.  
ACCESSION BG717067  
VERSION BG717067.1 GI:13996254  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

Query Match 90.0%; Score 18; DB 9; Length 516;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20  
|||||  
Db 125 AAAGAGCGCTCCAAAGCC 142

## RESULT 3

Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLNL0728 row: a column: 12  
 High quality sequence stop: 716.  
 Location/Qualifiers  
 1. .891  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4821395"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC 97"  
 /notes="Organ: testis; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
 (gtcgag); Oligo-dT primed using primer  
 5'-TTTATTTTATTTTATTTVN-3', size-selected for average  
 insert size 2.2 kb and normalized to ROT 5. This is a  
 primary library enriched for full-length clones and  
 constructed using the Cap-trapper method (Carninci, in  
 preparation). Library constructed by M. Brownstein  
 (NIMH/NHGRI, National Institutes of Health). Note: this is  
 a NIH\_MGC Library."

## ORIGIN

Query Match 90.0%; Score 18; DB 12; Length 891;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAGAGCGCTCCAAAGCC 20

Db 425 AAAGAGCGCTCCAAAGCC 408

RESULT 5  
 BX223249  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

EX223249 597 bp DNA linear GSS 29-JAN-2003  
 Danio rerio genomic clone DKEX-50L13, genomic survey sequence.  
 BX223249  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

## FEATURES

source  
 1. .597  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEX-50L13"  
 /tissue\_type="Testis"  
 /note="vector pindigoBAC-536"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 597;  
 Best Local Similarity 94.7%; Pred. No. 4.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  
 Db

RESULT 6  
 CA354389/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

## FEATURES

source  
 1. .615  
 /organism="Oncorhynchus mykiss"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8022"  
 /clone="IR77106\_C\_E03"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="NCCCWCA 1RT"  
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
 library made from pooled tissue from brain, gill, liver,  
 spleen, muscle, and kidney."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 615;  
 Best Local Similarity 94.7%; Pred. No. 4.3e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20

Db 119 TAGAGAGCGCTCCAAAGCC 101

RESULT 7  
 AQ756548  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS

## TITLE

Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome

2 TAAAGAGCGCTCCAAAGCC 20  
 |||||  
 35 TAAAGAGCGCTCCAAAGCC 53  
 |||||  
 CA354389 615 bp mRNA linear EST 05-NOV-2002  
 626102 NCCWCA 1RT Oncorhynchus mykiss cDNA clone IR77106\_C\_E03 5',  
 mRNA sequence.  
 CA354389  
 CA354389.1 GI:24599576  
 EST.  
 Oncorhynchus mykiss (rainbow trout)  
 Oncorhynchus mykiss  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 1 (bases 1 to 615)  
 Rexroad, C.B. and Keele, J.W.  
 Sequence analysis of a rainbow trout normalized cDNA library  
 Unpublished (2002)  
 Contact: Rexroad CE  
 USDA, ARS, National Center for Cool and Cold Water Aquaculture  
 11876 Leetown Road, Kearneysville, WV 25430, USA  
 Tel: 304 724 8340 x2129  
 Fax: 304 725 0351  
 Email: crexroad@nccwca.ars.usda.gov  
 Single pass sequencing. Bases called with phred v0.020425.c and  
 trimmed with the aid of the trim\_alt option. Vector identified by  
 cross match v0.990329.  
 Seq primer: AGCGGATACAAATTTCACACAGGA.

## FEATURES

Location/Qualifiers  
 1. .615  
 /organism="Oncorhynchus mykiss"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8022"  
 /clone="IR77106\_C\_E03"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="NCCCWCA 1RT"  
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
 library made from pooled tissue from brain, gill, liver,  
 spleen, muscle, and kidney."

Query Match 87.0%; Score 17.4; DB 14; Length 615;  
 Best Local Similarity 94.7%; Pred. No. 4.3e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAAAGAGCGCTCCAAAGCC 20

Db 119 TAGAGAGCGCTCCAAAGCC 101

RESULT 7  
 AQ756548 631 bp DNA linear GSS 27-JUL-1999  
 HS\_5373\_B2\_B03\_T7A RPCI-11 Human Male BAC Library Homo sapiens  
 genomic clone Plate=949 Col=6 Row=D, genomic survey sequence.  
 AQ756548  
 AQ756548.1 GI:5621206  
 GSS.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 631)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
 Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
9380589  
1049764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm)) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 949 row: D column: 6

Seq primer: T7

Class: BAC ends

High quality sequence stop: 631.

Location/Qualifiers

1. .631  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=949 Col=6 Row=D"  
/sex="male"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/notes="Vector: pBAC3.6, Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

# ORIGIN

Query Match 85.0%; Score 17; DB 28; Length 631;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TAAAGAGCGTCCAAAG 18

Db 548 TAAAGAGCGTCCAAAG 564

RESULT 8  
CF835325/c  
LOCUS  
DEFINITION  
UCRCS03.01A20\_r Washington Navel Orange Shoot Meristem cDNA Library  
Citrus sinensis cDNA clone CS\_PBA01A20, mRNA sequence.

CF835325

CF835325.1 GI:38050875

EST.

Source

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

Citrus sinensis

/mol\_type="mRNA"  
/cultivar="Parent Washington Navel"  
/db\_xref="taxon:2711"  
/clone="CS\_PBA01A20"  
/tissue\_type="Shoot meristem"  
/dev\_stage="10 year old trees"  
/lab\_host="E. coli TUC121"  
/clone\_lib="Washington Navel Orange Shoot Meristem cDNA Library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
Site 1: EcoRI; Site 2: XhoI; Parent Washington Navel Orange trees on Troyer rootstock (UCR 16K) were the source of tissue. Trees, at UC Riverside Agricultural Operations, were planted October 12, 1992. In each of 17 reps one tree on Troyer rootstock was initially treated with Enzone, one with Alliette and Nemacur, and one was left untreated. These treatments were discontinued in 1998. At the time of sampling, there were differences in the apparent health and size of the trees on Troyer rootstock. Fall-flush shoots were sampled in early November 2002 to minimize the number of floral shoot meristems. Federici and Mu (Roose lab) harvested meristems only from trees that appeared to be healthy and had a large number of young shoot tips on the day of collection. The average weight of a meristem was about 2 mg. Federici noted that there were quite a few insects and signs of insect damage to the shoot tips. Mealy bugs, thrips and aphids were observed, plus a few very tiny fast moving insects that may have been mites or crawler stage of scale (although Federici did not see any mature scale). It was not difficult to avoid collecting most of these because they were easy to see with the dissecting microscope. It was harder to exclude the frass. Some frass was definitely retained in the samples. Missues were snap frozen and then stored at -80C until further processing. Fenton (Close lab) purified RNA by the phenol method described in J. Japanese Soc. Hort. Sci. 1996. 64 (4): 809-814, purified poly(A) mRNA using a PolyATrack mRNA Isolation System IV (Promega), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised one million pfu from the primary library to produce a phagemid population. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the Arizona Genomics Institute, University of Arizona (Kim, Kudrna, Wing, Yu). Chromatogram files were downloaded by FTP to UC Riverside (by Close), then processed at UC Riverside (by Wanmaker, Close lab) using the HarVest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

# ORIGIN

Query Match 85.0%; Score 17; DB 14; Length 685;  
Best Local Similarity 100.0%; Pred. No. 7.1e-02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TAAAGAGCGTCCAAAG 18

Db 313 TAAAGAGCGTCCAAAG 297

RESULT 9  
BB529596/c  
LOCUS  
DEFINITION  
BB529596 RIKEN full-length enriched, 0 day neonate lung Mus musculus cDNA clone E030003D02.3, mRNA sequence.

BB529596

BB529596.1 GI:9581054

EST.

Mus musculus (house mouse)

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 277)

Yamamoto, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirazawa, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuro, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Yamamatsu, M. and Hayashizaki, Y.

RIKEN Mouse SSTS (Konno, H., et al.)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9216

Fax: 81-45-503-9216

URL: <http://genome.res.gsc.riken.go.jp/>

Email: [genome-res@gsc.riken.go.jp](mailto:genome-res@gsc.riken.go.jp)

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S., Saeki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotranscription and thermostabilization of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, Y., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Location/Qualifiers

1..277

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="E030003D02"

/tissue\_type="lung"

/dev\_stage="0 day neonate"

/lab\_host="DH10B"

/clone\_lib="RIKEN full-length enriched, 0 day neonate lung"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGCGCCGCACTCGAGTTTGTGTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTTATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 277;

Best Local Similarity 90.0%; Pred. No. 6.8e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20

Db 246 TCAAGAGCACTCCAAAGCC 227

RESULT 10

LOCUS

BE022711/c 331 bp mRNA linear EST 03-DEC-2001  
 sm876303.y1 Gm-cl015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl015-7061 5' similar to TR:041442 Q41442 ALPHA-AMYLASE  
 PRECURSOR. i, mRNA sequence.

ACCESSION

VERSION BE022711.1 GI:8285152

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM Glycine max

DEFINITION Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 331)

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCam, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: [ccu@resgen.com](mailto:ccu@resgen.com)

Insert length: 551 Std Error: 0.00

Seq primer: -40RP from Gibco.

FEATURES

Location/Qualifiers

1..331

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-7061"

/tissue\_type="Mature flowers, field grown plants"

/lab\_host="XL10-Gold"

/clone\_lib="Gm-cl015"

/note="Vector: pBluescript II XE; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 331;

Best Local Similarity 90.0%; Pred. No. 7.2e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20

Db 246 TCAAGAGCACTCCAAAGCC 227



REFERENCE 1 (bases 1 to 425)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Expelting,J., Corryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Public Soybean EST Project  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 High quality sequence stop: 416.

FEATURES  
 source  
 1..425  
 /location/Qualifiers  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl073-1275"  
 /tissue type="seedlings induced for symptoms of SDS  
 (Sudden Death Syndrome) disease"  
 /dev stage="2-3 weeks old"  
 /lab\_hosts="DH10B"  
 /clone\_lib="Gm-cl073"  
 /note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from 2-3 week old seedlings that were induced for symptoms  
 of SDS (Sudden Death Syndrome) disease by the  
 translocation of culture filtrate of *Fusarium solani* f.  
 sp. glycinis (Plant Cell Report 18:375-380). Cultivar  
 Williams 82 is susceptible to the disease SDS. Plant  
 tissue (expanded leaves, folded leaves, and new shoots)  
 were collected at 1, 6, 24, and 48 hrs. after inoculation  
 and their mRNA pooled equally for cDNA construction. The  
 library was prepared using the Stratagene pBluescript II  
 SK(+) library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(GT) sequence with an XhoI restriction site. EcoRI  
 adaptors were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA insert is protected  
 from XhoI digestion via methylation during first strand  
 synthesis. The cDNA fragments were directionally cloned  
 into the EcoRI-XhoI restriction site of the pBluescript  
 vector. The ligated cDNA fragments were transformed into  
 E.coli ElectroMax DH10B host cells. Plants were inoculated  
 by Shuxian Li (Glen Hartman lab, University of Illinois).  
 Library was constructed by Reena Philip and Steve Clough  
 (Lila Vodkin lab, University of Illinois)."

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 12; Length 425;  
 Best Local Similarity 90.0%; Pred. No. 7.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20  
 |||||  
 Db 313 TTAAGAGCGCTTCCTCAAGCC 294

RESULT 14  
 A1228821  
 LOCUS EST225516  
 DEFINITION Normalized rat brain, Bento Soares Rattus sp. cDNA clone  
 RRC293 3' end, mRNA sequence.  
 ACCESSION A1228821  
 VERSION A1228821.1 GI:4135385

QY 1 TTAAGAGCGCTCCAAAGCC 20  
 |||||  
 Db 313 TTAAGAGCGCTTCCTCAAGCC 294

RESULT 14  
 A1228821  
 LOCUS EST225516  
 DEFINITION Normalized rat brain, Bento Soares Rattus sp. cDNA clone  
 RRC293 3' end, mRNA sequence.  
 ACCESSION A1228821  
 VERSION A1228821.1 GI:4135385

KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 427)  
 AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Karlavage,A.R. and Adams,M.D.  
 TITLE Rat genome Project: Generation of a Rat EST (RSET) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT On Oct 30, 1998 this sequence version replaced gi:3812708.  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 Location/Qualifiers  
 1..427  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 /clone="RRC293"  
 /clone\_lib="Normalized rat brain, Bento Soares"  
 /note="Organ: brain; Vector: pT73Pac; Site\_1: EcoRI;  
 Site\_2: NotI"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 427;  
 Best Local Similarity 90.0%; Pred. No. 7.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAAGAGCGCTCCAAAGCC 20  
 |||||  
 Db 182 TCAAGAGCGCTTCCTCAAGCC 201

RESULT 15  
 CF863082  
 LOCUS CF863082  
 DEFINITION PF863082 497 bp mRNA linear EST 31-OCT-2003  
 PSZ005XN13f USDA-IFAPs:Expression of Phytophthora sojae genes  
 during infection and propagation\_sds Phytophthora sojae cDNA clone  
 szs005N13 5, mRNA sequence.  
 ACCESSION CF863082  
 VERSION CF863082.1 GI:38117708  
 KEYWORDS EST.  
 SOURCE Phytophthora sojae  
 ORGANISM Phytophthora sojae  
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
 Phytophthora.  
 REFERENCE 1 (bases 1 to 497)  
 AUTHORS Tyler,B.  
 TITLE Tyler,B. Not Published  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Tyler B  
 Tyler lab  
 VBI  
 1880 Pratt Dr., Blacksburg, VA 24061, USA  
 Tel: 540-231-7318  
 Email: bmtyley@vt.edu  
 PCR Primers  
 FORWARD: BK reverse primer  
 BACKWARD: BK reverse primer  
 Plate: 005 row: N column: 13  
 Seq primer: BK reverse primer  
 High quality sequence stop: 497.  
 Location/Qualifiers  
 1..497  
 /organism="Phytophthora sojae"  
 /mol\_type="mRNA"

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 785.506 Seconds  
(without alignments)  
434.641 Million cell updates/sec

Title: US-09-877-819B-37

Perfect score: 26

Sequence: 1 tcaactatgcgcggtttgtacagac 26

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.man.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	26	100.0	244	9	AF098794	Homo sapi
2	26	100.0	246	9	AF346471	Homo sapi
3	26	100.0	246	9	HS0103X2	H. sapiens M
4	26	100.0	246	9	HS0201X2	H. sapiens M
5	26	100.0	246	9	HS0201X2	H. sapiens M
6	26	100.0	246	9	HS0201X2	H. sapiens M
7	26	100.0	252	9	HS248473	H. sapiens H
8	26	100.0	257	9	AF165160	Homo sapi
9	26	100.0	258	9	AF118120	Homo sapi
10	26	100.0	267	6	AX237167	Sequence
11	26	100.0	267	6	AX237352	Sequence
12	26	100.0	268	9	HS0103X2	H. sapiens H
13	26	100.0	272	6	AX237066	Sequence
14	26	100.0	273	9	AF076284	Homo sapi
15	26	100.0	279	9	AF076285	Homo sapi
16	26	100.0	286	9	HS075556	Homo sapi
17	26	100.0	287	9	AF015295	Homo sapi
18	26	100.0	294	6	AX237304	Sequence
19	26	100.0	294	6	AX237554	Sequence
20	26	100.0	396	6	BD058253	Secreted
21	26	100.0	466	6	BD058339	Secreted
22	26	100.0	476	6	AX884252	Sequence
23	26	100.0	476	6	BD023862	Sequence
24	26	100.0	576	6	AX884251	Sequence
25	26	100.0	576	6	BD023861	Sequence
26	26	100.0	661	6	E00485	DNA sequenc
27	26	100.0	661	6	I03086	Sequence 5
28	26	100.0	690	9	I03086	Sequence 5
29	26	100.0	818	9	HUMHDFPA	Homo sapien
30	26	100.0	1048	9	HSSBA1	Human mRNA
31	26	100.0	1140	6	I03088	Sequence 7
32	26	100.0	1201	6	E00484	DNA sequenc
33	26	100.0	1201	6	I03006	Sequence 6
34	26	100.0	1201	9	I03423	Sequence 6
35	26	100.0	1201	9	BC009956	Homo sapi
36	26	100.0	1259	6	AX552229	Sequence
37	26	100.0	1480	6	AX780118	Sequence
38	26	100.0	1480	6	AX780119	Sequence
39	26	100.0	2986	9	HUMHDFC02	Human MEC C
40	26	100.0	14646	9	HS0103X2	Human HLA-S
41	26	100.0	64380	9	EX120009	Human DNA
42	26	100.0	106728	9	AL805913	Human DNA
43	26	100.0	124899	9	AL645931	Human DNA
44	26	100.0	181228	2	AC011086	Human DNA
45	22.8	87.7	214	9	F000000000	Human DNA
46	22.8	87.7	265	9	AF092049	Homo sapi
47	22.8	87.7	326	9	AF013767	Homo sapi
48	21.2	81.5	349980	6	AX344560	Sequence
49	20.8	75.4	1834	9	HUMHDFC02	Human MEC C
50	19.6	75.4	22847	9	HS0103X2	Human HLA-S
51	19.6	75.4	93842	9	AL845446	Human DNA
52	19.6	75.4	158033	9	AL645940	Human DNA
53	19.6	75.4	175737	9	HS0103X2	Human DNA
54	19.6	75.4	349980	6	AX344560	Sequence
55	18.6	71.5	2071	5	AY391467	Danio rer
56	18.6	71.5	2099	5	BC044510	Danio rer
57	18.6	71.5	164325	10	AL845258	Mouse DNA
58	18.6	71.5	170397	9	AC119677	Homo sapi
59	18.6	71.5	177792	9	CNS06C7M	Human chr
60	18.6	71.5	229220	10	AC091519	Mus muscu
61	18.6	71.5	239550	1	AP001511	Bacillus
62	18.2	70.0	68487	2	AC109287	Continuation (4 of
63	18.2	70.0	74658	8	OS24243961	Oryza sat
64	18.2	70.0	162581	2	AC121658	Rattus no

C 66	18.2	70.0	164805	8	OSUN00174	AL662970	Oryza sat	139	17.6	67.7	545	8	AY437764	AY437764	Sitococcu
C 67	18.2	70.0	193311	2	BD118694	AC118694	Mus muscu	140	17.6	67.7	545	8	AY437765	AY437765	Sitococcu
C 68	18	69.2	18	6	BD104220	BD104220	Kit and m	141	17.6	67.7	545	8	AY437766	AY437766	Sitococcu
C 69	18	69.2	1263	10	AB084158	AB084158	Cavia por	142	17.6	67.7	545	8	AY437767	AY437767	Sitococcu
C 70	18	69.2	2136	3	AY071523	AY071523	Drosophi	143	17.6	67.7	545	8	AY437768	AY437768	Sitococcu
C 71	18	69.2	2565	9	HUMH8XA	MY1591	Human MHC c	144	17.6	67.7	545	8	AY437769	AY437769	Sitococcu
C 72	18	69.2	44117	2	AC017312	AC017312	Drosophi	145	17.6	67.7	545	8	AY437770	AY437770	Sitococcu
C 73	18	69.2	57160	2	AC116457	AC116457	Homo sapi	146	17.6	67.7	641	8	AF260816	AF260816	Sitococcu
C 74	18	69.2	67375	2	AC104767	AC104767	Homo sapi	147	17.6	67.7	12989	1	AE000449	AE000449	Sitococcu
C 75	18	69.2	68444	2	AC114586	AC114586	Homo sapi	148	17.6	67.7	64166	2	AC012728	AC012728	Escherich
C 76	18	69.2	83376	9	AP004709	AP004709	Homo sapi	149	17.6	67.7	95272	3	AC005298	AC005298	Drosophi
C 77	18	69.2	87684	4	AY152828	AY152828	Felis cat	150	17.6	67.7	107770	5	EX323010	EX323010	Zebrafish
C 78	18	69.2	101882	9	AC021089	AC021089	Homo sapi	151	17.6	67.7	136393	2	AC1136952	AC1136952	Danio rer
C 79	18	69.2	107025	2	AL139235 <sup>3</sup>	Continuation (4 of		152	17.6	67.7	142237	2	AC108180	AC108180	Felis cat
C 80	18	69.2	110000	2	LMFLCHR35 <sup>23</sup>	Continuation (24 of		153	17.6	67.7	143885	2	EX890544	EX890544	Danio rer
C 81	18	69.2	132445	4	AY152826	Pellis cat		154	17.6	67.7	158601	2	EX842692	EX842692	Danio rer
C 82	18	69.2	140207	9	HSL1000E10	AL096773	Human DNA	155	17.6	67.7	162835	10	AC121866	AC121866	Mus muscu
C 83	18	69.2	149001	9	AC107909	AC107909	Homo sapi	156	17.6	67.7	164508	2	AC104130	AC104130	Homo sapi
C 84	18	69.2	157034	2	AC125237	AC125237	Homo sapi	157	17.6	67.7	169757	5	AL627256	AL627256	Zebrafish
C 85	18	69.2	160763	9	AC097482	AC097482	Homo sapi	158	17.6	67.7	170727	3	AC104514	AC104514	Drosophi
C 86	18	69.2	161803	9	AC021028	AC021028	Homo sapi	159	17.6	67.7	178238	4	AC092519	AC092519	Felis cat
C 87	18	69.2	164884	3	AC011063	AC011063	Drosophi	160	17.6	67.7	183946	2	AC009164	AC009164	Homo sapi
C 88	18	69.2	167834	9	AC108195	AC105195	Homo sapi	161	17.6	67.7	189341	10	AC117671	AC117671	Mus muscu
C 89	18	69.2	180327	2	AC080110	AC080110	Homo sapi	162	17.6	67.7	198949	9	AC010287	AC010287	Homo sapi
C 90	18	69.2	180236	9	AP004294	AP004294	Homo sapi	163	17.6	67.7	199287	2	AC068659	AC068659	Homo sapi
C 91	18	69.2	181948	2	AC079008	AC079008	Homo sapi	164	17.6	67.7	201758	2	AC102358	AC102358	Mus muscu
C 92	18	69.2	187385	9	CNS01DU9	AL133224	Human chr	165	17.6	67.7	208913	2	AC011240	AC011240	Homo sapi
C 93	18	69.2	193924	3	AC009910	AC009910	Drosophi	166	17.6	67.7	212349	9	AC008009	AC008009	Homo sapi
C 94	18	69.2	249233	2	AC116670	AC116670	Mus muscu	167	17.6	67.7	212737	2	AC147207	AC147207	Silurana
C 95	18	69.2	260249	3	AB03610	AE003610	Drosophi	168	17.6	67.7	218226	2	AC126634	AC126634	Rattus no
C 96	18	69.2	261474	2	AC094196	AC094196	Rattus no	169	17.6	67.7	218986	2	AC111652	AC111652	Rattus no
C 97	17.8	68.5	160654	2	AC011879	AC011879	Homo sapi	170	17.6	67.7	219417	9	AC096826	AC096826	Rattus no
C 98	17.8	68.5	168473	9	AL512324	AL512324	Human DNA	171	17.6	67.7	220173	9	AC012183	AC012183	Homo sapi
C 99	17.8	68.5	168473	9	AC016568	AC016568	Homo sapi	172	17.6	67.7	22045	2	AC023354	AC023354	Mus muscu
C 100	17.6	67.7	30	6	AY546653	AY546653	Sequence	173	17.6	67.7	223281	2	AC125010	AC125010	Mus muscu
C 101	17.6	67.7	543	8	AY163798	AY163798	Sitococcu	174	17.6	67.7	225379	2	AC117693	AC117693	Mus muscu
C 102	17.6	67.7	544	8	AY163799	AY163799	Sitococcu	175	17.6	67.7	240732	2	AC133848	AC133848	Rattus no
C 103	17.6	67.7	544	8	AY163799	AY163799	Sitococcu	176	17.6	67.7	243754	2	AC146822	AC146822	Otolemur
C 104	17.6	67.7	544	8	AY163799	AY163799	Sitococcu	177	17.6	67.7	247187	2	AC094928	AC094928	Rattus no
C 105	17.6	67.7	544	8	AY168970	AY168970	Sitococcu	178	17.6	67.7	254908	3	AE003801	AE003801	Drosophi
C 106	17.6	67.7	544	8	AY437740	AY437740	Sitococcu	179	17.6	67.7	271669	2	AC094772	AC094772	Rattus no
C 107	17.6	67.7	544	8	AY437741	AY437741	Sitococcu	180	17.6	67.7	288925	2	AC146422	AC146422	Par trogl
C 108	17.6	67.7	544	8	AY437742	AY437742	Sitococcu								
C 109	17.6	67.7	544	8	AY437743	AY437743	Sitococcu								
C 110	17.6	67.7	544	8	AY437746	AY437746	Sitococcu								
C 111	17.6	67.7	544	8	AY437748	AY437748	Sitococcu								
C 112	17.6	67.7	544	8	AY437757	AY437757	Sitococcu								
C 113	17.6	67.7	544	8	AY437761	AY437761	Sitococcu								
C 114	17.6	67.7	544	8	AY437762	AY437762	Sitococcu								
C 115	17.6	67.7	544	8	AY437763	AY437763	Sitococcu								
C 116	17.6	67.7	544	8	AY437766	AY437766	Sitococcu								
C 117	17.6	67.7	544	8	AY437767	AY437767	Sitococcu								
C 118	17.6	67.7	544	8	AY437771	AY437771	Sitococcu								
C 119	17.6	67.7	544	8	AY437773	AY437773	Sitococcu								
C 120	17.6	67.7	545	8	AY163791	AY163791	Sitococcu								
C 121	17.6	67.7	545	8	AY163792	AY163792	Sitococcu								
C 122	17.6	67.7	545	8	AY163794	AY163794	Sitococcu								
C 123	17.6	67.7	545	8	AY163796	AY163796	Sitococcu								
C 124	17.6	67.7	545	8	AY163797	AY163797	Sitococcu								
C 125	17.6	67.7	545	8	AY168969	AY168969	Sitococcu								
C 126	17.6	67.7	545	8	AY437738	AY437738	Sitococcu								
C 127	17.6	67.7	545	8	AY437739	AY437739	Sitococcu								
C 128	17.6	67.7	545	8	AY437742	AY437742	Sitococcu								
C 129	17.6	67.7	545	8	AY437744	AY437744	Sitococcu								
C 130	17.6	67.7	545	8	AY437745	AY437745	Sitococcu								
C 131	17.6	67.7	545	8	AY437747	AY437747	Sitococcu								
C 132	17.6	67.7	545	8	AY437749	AY437749	Sitococcu								
C 133	17.6	67.7	545	8	AY437751	AY437751	Sitococcu								
C 134	17.6	67.7	545	8	AY437752	AY437752	Sitococcu								
C 135	17.6	67.7	545	8	AY437753	AY437753	Sitococcu								
C 136	17.6	67.7	545	8	AY437756	AY437756	Sitococcu								
C 137	17.6	67.7	545	8	AY437758	AY437758	Sitococcu								
C 138	17.6	67.7	545	8	AY437759	AY437759	Sitococcu								
						AY437760	Sitococcu								

## ALIGNMENTS

RESULT 1	AF098794	244 bp	DNA	linear	PRI	01-SEP-2000
LOCUS	Homo sapiens MHC class II antigen (HLA-DPA1) gene, exon 2 and partial cds.					
DEFINITION	AF098794					
ACCESSION	AF098794.1	GI:3859561				
VERSION	AF098794					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	McTernan, C.H., Miljovic, C.H., Cockram, C.S. and Barnett, A.H.					
AUTHORS	McTernan, C.H., Miljovic, C.H., Cockram, C.S. and Barnett, A.H.					
TITLE	The nucleotide sequence of two new DP alleles, DPA1*02015 and DPA1*8401, identified in a Chinese subject					
JOURNAL	Tissue Antigens 56 (1), 95-98 (2000)					
MEDLINE	20412568					
PUBMED	10959363					
REFERENCE	2 (bases 1 to 244)					
AUTHORS	Perry, C.L., Mijovic, C.H., Cockram, C.S. and Barnett, A.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (14-OCT-1998) Medicine, Birmingham University, Clinical Research Block, Queen Elizabeth Hospital, Edgbaston, Birmingham B-5 2TH, England					
FEATURES	Location/Qualifiers					

```

source
1. .244
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.1"
<1. .>244
/gene="HLA-DPA1"
<1. .>244
/gene="HLA-DPA1"
<1. .>244
/product="MHC class II antigen"
/codon_start=1
/protein_id="AAC72845.1"
/db_xref="GI:3859562"
/translation="DFVSTYAAAFVQTHRPTGHEFMFEFDDDDMEFYVDLKKETVWHLEE
FGRAFSFPAQGGLANIAILNNLNLTIQRSNHTQATN"
<1. .244
/gene="HLA-DPA1"
/number=2

exon
Query Match 100.0%; Score 26; DB 9; Length 244;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

QY 1 TCAACTTATGCCGCGTTGTACAGAC 26
|||||
DB 10 TCAACTTATGCCGCGTTGTACAGAC 35

RESULT 2
AF346471
LOCUS Homo sapiens MHC class II antigen (HLA-DPA1) gene, HLA-DPA1-new
DEFINITION
ACCESSION AF346471
VERSION AF346471.1 GI:13448661
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
Grais, S.E., Begovich, A. and Mangano, J.
Unpublished
JOURNAL
2 (bases 1 to 246)
Grais, S.E., Begovich, A. and Mangano, J.
Direct Submission
AUTHORS
Submitted (07-FEB-2001) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
JOURNAL
1145 Atlantic Ave., Alameda, CA 94501, USA
FEATURES
Location/Qualifiers
1. .246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p"
<1. .>246
/gene="HLA-DPA1"
/allele="new"
<1. .>246
/product="MHC class II antigen"
/codon_start=3
<1. .>246
/gene="HLA-DPA1"
/number=2

gene
Query Match 100.0%; Score 26; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

QY 1 TCAACTTATGCCGCGTTGTACAGAC 26
|||||
DB 12 TCAACTTATGCCGCGTTGTACAGAC 37

RESULT 4
HS0201X2
LOCUS H.sapiens MHC class II HLA-DPA1*0201 gene (exon 2)
DEFINITION
ACCESSION HS0201X2
VERSION X82394.1 GI:572739

```

```

/translation="DFVSTYAAAFVQTHRPTGHEFMFEFDDDDMEFYVDLKKETVWHLEE
FGRAFSFPAQGGLANIAILNNLNLTIQRSNHTQATN"
1. .246
/gene="HLA-DPA1"
/number=2

exon
Query Match 100.0%; Score 26; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

QY 1 TCAACTTATGCCGCGTTGTACAGAC 26
|||||
DB 12 TCAACTTATGCCGCGTTGTACAGAC 37

RESULT 3
HS0103X2
LOCUS H.sapiens MHC class II HLA-DPA1*0103 gene (exon 2)
DEFINITION
ACCESSION HS0103X2
VERSION X82390.1 GI:565028
KEYWORDS MHC class II HLA-DPA1.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
Rozenmuller, E.H., Bouwens, A.G., van Oort, E., Versluis, L.F.,
Marsh, S.G., Bodmer, J.G. and Tilanus, M.G.
Sequencing-based typing reveals new insight in HLA-DPA1
Polymorphism
JOURNAL
Tissue Antigens 45 (1), 57-62 (1995)
MEDLINE
95242313
PUBMED
7725312
REFERENCE
2 (bases 1 to 246)
Rozenmuller, E.H.
Direct Submission
AUTHORS
Submitted (19-OCT-1994) E.H. Rozenmuller, Diagnostic DNA Laboratory,
University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA
Utrecht, NETHERLANDS
COMMENT
Related sequences: S52453 and D14344.
Related sequences: S52453 and D14344.
FEATURES
Location/Qualifiers
1. .246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/haplotype="DPA1*0103"
/cell_line="LB and LG2"
1. .246
/gene="MHC class II HLA-DPA1 gene"
1. .246
/gene="MHC class II HLA-DPA1 gene"
/number=2

gene
Query Match 100.0%; Score 26; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

QY 1 TCAACTTATGCCGCGTTGTACAGAC 26
|||||
DB 12 TCAACTTATGCCGCGTTGTACAGAC 37

RESULT 4
HS0201X2
LOCUS H.sapiens MHC class II HLA-DPA1*0201 gene (exon 2)
DEFINITION
ACCESSION HS0201X2
VERSION X82394.1 GI:572739

```

KEYWORDS MHC class II HLA-DPA1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 246)  
Rozenmuller,E.H., Bouwers,A.G., van Oort,E., Versluis,L.F., Marsh,S.G., Bodmer,J.G. and Tilianus,M.G.

TITLE Sequencing-based typing reveals new insight in HLA-DPA1 polymorphism

JOURNAL Tissue Antigens 45 (1), 57-62 (1995)  
MEDLINE 95242313  
PUBMED 7725312

REFERENCE 2 (bases 1 to 246)  
Rozenmuller,E.H.  
AUTHORS Direct Submission  
JOURNAL Submitted (19-OCT-1994) E.H. Rozenmuller, Diagnostic DNA Laboratory, University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA Utrecht, NETHERLANDS

COMMENT Related sequences: S52453 and D14344.  
Related sequences: S52453, D14344 and M83906.

FEATURES  
source Location/Qualifiers  
1..246  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/haplotype="DPA1\*0201"  
/cell\_line="DAUDI and AKTBA"  
1..226  
/gene="MHC class II HLA-DPA1 gene"  
1..226  
/gene="MHC class II HLA-DPA1 gene"  
/number=2

gene  
exon

ORIGIN  
Query Match 100.0%; Score 26; DB 9; Length 246;  
Best Local Similarity 100.0%; Pred.No. 1.7;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCCGCGTTTGTACAGAC 26  
|||||  
12 TCAACTTATGCCGCGTTTGTACAGAC 37

Query Match 100.0%; Score 26; DB 9; Length 246;  
Best Local Similarity 100.0%; Pred.No. 1.7;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCCGCGTTTGTACAGAC 26  
|||||  
12 TCAACTTATGCCGCGTTTGTACAGAC 37

RESULT 5  
HSHLADPA1  
LOCUS H.sapiens HLA-DPA1 gene, exon 2.  
DEFINITION X78198  
ACCESSION X78198.1 GI:461355  
VERSION DPA1\*01new; HLA-DPA1 gene.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 246)  
Rozenmuller,E.H., Bouwers,A.G., van Oort,E., Versluis,L.F., Marsh,S.G., Bodmer,J.G. and Tilianus,M.G.

TITLE Sequencing-based typing reveals new insight in HLA-DPA1 polymorphism

JOURNAL Tissue Antigens 45 (1), 57-62 (1995)  
MEDLINE 95242313  
PUBMED 7725312

REFERENCE 2 (bases 1 to 246)  
Rozenmuller,E.H.  
AUTHORS Direct Submission  
JOURNAL Submitted (15-MAR-1994) E.H. Rozenmuller, Diagnostic DNA Lab, University Hospital Utrecht, G03.647, PO Box 85500, 3508 GA Utrecht, NETHERLANDS

FEATURES  
source Location/Qualifiers  
1..246  
/organism="Homo sapiens"

/mol\_type="genomic DNA"  
/isolate="SK"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/haplotype="DPA1\*01"  
1..246  
/gene="HLA-DPA1"  
1..246  
/gene="HLA-DPA1"  
/product="DPA1\*01new"  
/number=2

ORIGIN  
Query Match 100.0%; Score 26; DB 9; Length 246;  
Best Local Similarity 100.0%; Pred.No. 1.7;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCCGCGTTTGTACAGAC 26  
|||||  
12 TCAACTTATGCCGCGTTTGTACAGAC 37

RESULT 6  
HSZ48473  
LOCUS H.sapiens HLA-DPA1 gene for first domain of MHC class 2 molecule,  
DEFINITION alpha-chain (allele DPA1\*0203).  
ACCESSION 248473.1 GI:1770743  
VERSION alpha-chain; MHC class 2 molecule.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 252)  
Muntau,B., Thyte,T., Pirmez,C. and Horstmann,R.D.  
AUTHORS A novel DPA1 allele (DPA1\*0203) composed of known epitopes  
JOURNAL Tissue Antigens 49 (5), 668-669 (1997)  
MEDLINE 97378898  
PUBMED 9234495

REFERENCE 2 (bases 1 to 252)  
Muntau,B.  
AUTHORS Direct Submission  
JOURNAL Submitted (23-FEB-1995) Birgit Muntau, Molecular Genetics, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Bernhard-Nocht-Str.74, 20359 Hamburg, Germany  
COMMENT On Jan 9, 1997 this sequence version replaced gi:683569.  
FEATURES  
source Location/Qualifiers  
1..252  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="Brazilian White"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="TC 48-3"  
/sex="female"  
/tissue\_type="blood"  
/dev\_stage="adult"  
31..252  
/gene="HLA-DPA1"  
31..252  
/gene="HLA-DPA1"  
/product="first domain of MHC class 2 molecule, alpha-chain"  
/note="allele DPA1\*0203"  
/number=2

gene  
exon

ORIGIN  
Query Match 100.0%; Score 26; DB 9; Length 252;  
Best Local Similarity 100.0%; Pred.No. 1.7;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCCGCGTTTGTACAGAC 26





```

RESULT 10
AX237352
LOCUS          267 bp      DNA
DEFINITION    Sequence 328 from Patent WO0164886.
ACCESSION    AX237352
VERSION      AX237352.1  GI:15796906
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Gaiger, A., Algate, P.A. and Mannion, J.
TITLE        Compositions and methods for the detection, diagnosis and therapy
              of hematological malignancies
JOURNAL      Patent: WO 0164886-A 328 07-SEP-2001;
              CORIXA CORPORATION (US)
FEATURES     source
              1..267
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 26; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 140 TCAACTTATGCGCGTTGTACAGAC 165

RESULT 11
HSHLADPAX
LOCUS          268 bp      DNA
DEFINITION    H.sapiens HLA-DPA1 gene.
ACCESSION    X83610
VERSION      X83610.1  GI:987073
KEYWORDS      HLA-DPA1 gene.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Versluis, L.F., Verdryn, W., Van der Zwan, A., Oudshoorn, M. and
              Tilanus, M.G.J.
TITLE        An update of the exon 2 sequence of the HLA-DPA1*02012 allele
JOURNAL      Tissue Antigens 46 (3 Pt 1), 206-207 (1995)
MEDLINE      96097411
PUBMED       8525481
REFERENCE    2 (bases 1 to 268)
AUTHORS      Tilanus, M.G.J.
TITLE        Direct Submission
JOURNAL      Submitted (20-DEC-1994) M.G.J. Tilanus, Diagnostic DNA Laboratory,
              Academic Hospital Utrecht, Heidelberglaan 100, PO Box 85500, 3508
              GA Utrecht, NETHERLANDS
COMMENT      Related sequence: L31624.
FEATURES     Location/Qualifiers
              1..268
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="6"
              /haplotype="DPA1*02012"
              <1..18
              19..264
              /gene="HLA-DPA1"
              19..264
              /gene="HLA-DPA1"
              265..>268
              intron
              gene
              exon
              intron
ORIGIN

```

```

Query Match      100.0%; Score 26; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 30 TCAACTTATGCGCGTTGTACAGAC 55

RESULT 12
AX237066
LOCUS          272 bp      DNA
DEFINITION    Sequence 42 from Patent WO0164886.
ACCESSION    AX237066
VERSION      AX237066.1  GI:15796620
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Gaiger, A., Algate, P.A. and Mannion, J.
TITLE        Compositions and methods for the detection, diagnosis and therapy
              of hematological malignancies
JOURNAL      Patent: WO 0164886-A 42 07-SEP-2001;
              CORIXA CORPORATION (US)
FEATURES     Location/Qualifiers
              1..272
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 26; DB 6; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 145 TCAACTTATGCGCGTTGTACAGAC 170

RESULT 13
AF076284
LOCUS          279 bp      DNA
DEFINITION    Homo sapiens isolate 913 MHC class II antigen (HLA-DPA1) gene,
              partial cds.
ACCESSION    AF076284
VERSION      AF076284.1  GI:5381292
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 279)
AUTHORS      Varney, M.D., Gavrilidis, A. and Abbott, W.
TITLE        DPA1 Polymorphism in Polynesians
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 279)
AUTHORS      Varney, M.D., Gavrilidis, A. and Abbott, W.
TITLE        Direct Submission
JOURNAL      Submitted (06-JUL-1998) Tissue Typing, Royal Melbourne Hospital,
              Grattan Street, Parkville, Vic 3050, Australia
FEATURES     Location/Qualifiers
              1..279
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /isolate="913"
              /db_xref="taxon:9606"
              /chromosome="6"
              <1..>279
              /gene="HLA-DPA1"
              <1..>279
              gene
              mRNA

```

```

/ gene="HLA-DPA1"
/ product="MHC class II antigen"
<1..>279
/ gene="HLA-DPA1"
/ codon_start=1
/ product="MHC class II antigen"
/ protein_id="AAD42927.1"
/ db_xref="GI:5381293"
/ translation="HVSFYAFVQTHRTGTFGEFDEDEQFYVDLDKKEVWHLSEF
GQTFSEPAQGLANLAIINNNLTIQRSNHTQATNGTPLYCLFLCSPT"
1..279
/ gene="HLA-DPA1"
/ number=2

ORIGIN
Query Match 100.0%; Score 26; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 7 TCAACTTATGCGCGTTGTACAGAC 32

RESULT 14
AF076285 279 bp DNA linear PRI 07-JUL-1999
LOCUS Homo sapiens isolate 63 MHC class II antigen (HLA-DPA1) gene,
DEFINITION partial cds.
ACCESSION AF076285
VERSION AF076285.1 GI:5381294
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Varney,M.D., Gavrilidis,A. and Abbott,W.
JOURNAL DPAL Polymorphism in Polynesians
REFERENCE 2 (bases 1 to 279)
AUTHORS Varney,M.D., Gavrilidis,A. and Abbott,W.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) Tissue Typing, Royal Melbourne Hospital,
Gratian Street, Parkville, Vic 3050, Australia
FEATURES
source
1..279
/ organism="Homo sapiens"
/ mol_type="genomic DNA"
/ isolate="63"
/ db_xref="taxon:9606"
/ chromosome="6"
<1..>279
/ gene="HLA-DPA1"
<1..>279
/ gene="HLA-DPA1"
/ product="MHC class II antigen"
<1..>279
/ gene="HLA-DPA1"
/ codon_start=1
/ product="MHC class II antigen"
/ protein_id="AAD42928.1"
/ db_xref="GI:5381295"
/ translation="HVSFYAFVQTHRTGTFGEFDEDEQFYVDLDKKEVWHLSEF
GRAPFSEPAQGLANLAIINNNLTIQRSNHTQANGTPLYCLFLCSPT"
1..279
/ gene="HLA-DPA1"
/ number=2

exon
Query Match 100.0%; Score 26; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

```

```

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 7 TCAACTTATGCGCGTTGTACAGAC 32

RESULT 15
HSU87556 286 bp DNA linear PRI 20-JAN-1998
LOCUS Homo sapiens MHC class II HLA-DPA1 antigen (HLA-D) gene, exon 2,
DEFINITION partial cds.
ACCESSION U87556
VERSION U87556.1 GI:2760313
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Steiner,L., Begovich,A. and Suraj,V.
JOURNAL Direct Submission
COMMENT Submitted (28-JAN-1997) Human Genetics, Roche Molecular Systems,
1145 Atlantic Ave., Alameda, CA 94501, USA
On Jan 8, 1998 this sequence version replaced gi:1842112.
FEATURES
source
1..286
/ organism="Homo sapiens"
/ mol_type="genomic DNA"
/ db_xref="taxon:9606"
/ chromosome="6"
<1..>286
/ gene="HLA-D"
<1..>261
/ gene="HLA-D"
/ note="MHC class II HLA-DPA1 antigen"
/ codon_start=3
/ protein_id="AAB97110.1"
/ db_xref="GI:2795772"
/ translation="DHVSFYAFVQTHRTGTFGEFDEDEQFYVDLDKKEVWHLSE
FGQAFSEPAQGLANLAIINNNLTIQRSNHTQATN"
16..261
/ gene="HLA-D"
/ number=2

exon
Query Match 100.0%; Score 26; DB 9; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 27 TCAACTTATGCGCGTTGTACAGAC 52

RESULT 16
AF015295 287 bp DNA linear PRI 29-SEP-1998
LOCUS Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*02013
DEFINITION allele); exon 2 and partial cds.
ACCESSION AF015295
VERSION AF015295.1 GI:3660653
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Steiner,L.L., Cavalli,A., Zimmerman,P.A., Boatín,B.A.,
Titanji,V.P., Bradley,J.E., Lucius,R., Nutman,T.B. and
Begovich,A.S.
JOURNAL Three new DP alleles identified in sub-Saharan Africa: DPB1*7401,
DPAL*02013, and DPAL*0302
Tissue Antigens 51 (6), 653-657 (1998)

```

```

MEDLINE 98357732
PUBMED 9694359
REFERENCE 2 (bases 1 to 287)
AUTHORS Steiner,L., Begovich,A. and Zimmerman,P.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1997) Human Genetics, Roche Molecular Systems,
Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA
FEATURES
source 1..287
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
<1..>287
/gene="HLA-DPA1"
/allele="HLA-DPA1*02013"
<1..15
/organism="HLA-DPA1"
/number=1
<16..>261
/organism="HLA-DPA1"
/product="MHC class II antigen"
<16..>261
/organism="HLA-DPA1"
/codon_start=3
/product="MHC class II antigen"
/protein_id="AAC61669.1"
/db_xref="GI:3660654"
/translations="DHVSTYAAAFVQTHPTGGERMFEDEQFVDLDKKEITVWHLLE
FGRAFPEAGGLANTALNNLTLQIRSNHTQAN"
16..261
/organism="HLA-DPA1"
/number=2
262..>287
/organism="HLA-DPA1"
/number=2
intron
Query Match 100.0%; Score 26; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 27 TCAACTTATGCGCGTTGTACAGAC 52

RESULT 17
AX237304/c
LOCUS AX237304 294 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 280 from Patent WO0164886.
ACCESSION AX237304
VERSION AX237304.1 GI:15796858
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gaiger,A., Algate,P.A. and Mannion,J.
TITLE Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A 280 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
intron
Query Match 100.0%; Score 26; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 27 TCAACTTATGCGCGTTGTACAGAC 52

RESULT 18
AX237554/c
LOCUS AX237554 294 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 530 from Patent WO0164886.
ACCESSION AX237554
VERSION AX237554.1 GI:15797108
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gaiger,A., Algate,P.A. and Mannion,J.
TITLE Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A 530 07-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..294
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
intron
Query Match 100.0%; Score 26; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 128 TCAACTTATGCGCGTTGTACAGAC 103

RESULT 19
BD058253
LOCUS BD058253 396 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (sESTs).
ACCESSION BD058253
VERSION BD058253.1 GI:22603859
KEYWORDS Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 396)
AUTHORS Jacobs,K., McCooy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
Treacy,M., Spaulding,V. and Agostino,M.J.
TITLE Secreted expressed sequence tags (sESTs)
JOURNAL Patent: JP 2001519666-A 108 23-OCT-2001;
GENETICS INSTITUTE INC
COMMENT PN JP 2001519666-A/108
PD 23-OCT-2001
PF 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..396
/organism="Zea mays"
/mol_type="genomic DNA"

```

```

/ab_xref="taxon:4577"

ORIGIN
Query Match          100.0%; Score 26; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TCAACTTATGCGCGGTTGTACAGAC 26
      |||||||
Db   163 TCAACTTATGCGCGGTTGTACAGAC 188

RESULT 20
BD058339          456 bp      DNA      linear      PAT 27-AUG-2002
LOCUS
DEFINITION
Secreted expressed sequence tags (seSTs).
ACCESSION
BD058339.1 GI:22603945
VERSION
JP 2001519666-A/194.
KEYWORDS
Zea mays
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
Clade; Pan.coideae; Andropogoneae; Zea.
REFERENCE
JACOBS,K., MCCOY,J.M., LAVALLIE,E.R., RACIE,L.A., MERBERG,D.,
TREACY,M., SPAULDING,V. and AGOSTINO,M.J.
Secreted expressed sequence tags (seSTs)
Patent: JP 2001519666-A 194 23-OCT-2001;
GENETICS INSTITUTE INC
COMMENT
FN  JP 2001519666-A/194
PD  23-OCT-2001
PF  10-APR-1998 JP 1998543068
PR  10-APR-1997 US  08/835913
PT  KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
PI  MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1..456
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match          100.0%; Score 26; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TCAACTTATGCGCGGTTGTACAGAC 26
      |||||||
Db   131 TCAACTTATGCGCGGTTGTACAGAC 156

RESULT 21
AX884252          476 bp      DNA      linear      PAT 18-DEC-2003
LOCUS
DEFINITION
Sequence 115 from Patent EP1033401.
ACCESSION
AX884252
VERSION
AX884252.1 GI:40039227
KEYWORDS
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1  Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
Expressed sequence tags and encoded human proteins
Patent: EP 1033401-A 115 06-SEP-2000;
GENSET (FR)

FEATURES
source
1..476
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          100.0%; Score 26; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TCAACTTATGCGCGGTTGTACAGAC 26
      |||||||
Db   206 TCAACTTATGCGCGGTTGTACAGAC 231

RESULT 22
BD023862          476 bp      DNA      linear      PAT 27-AUG-2002
LOCUS
DEFINITION
Sequence tag and encoded human, protein.
ACCESSION
BD023862
VERSION
BD023862.1 GI:22565085
KEYWORDS
JP 2001269182-A/108.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1  (bases 1 to 476)
AUTHORS
Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE
Sequence tag and encoded human protein
JOURNAL
Patent: JP 2001269182-A 108 02-OCT-2001;
GENSET

OS  Homo sapiens (human)
PN  JP 2001269182-A/108
PD  02-OCT-2001
PF  24-FEB-2000 JP 2000118773
PR  26-FEB-1999 US  60/122487
PI  JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PC  C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC  C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC score 10.1
FH Key Location/Qualifiers
FT CDS
FT sig_peptide 95..475
FT sig_peptide 95..187.

FEATURES
source
1..476
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          100.0%; Score 26; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TCAACTTATGCGCGGTTGTACAGAC 26
      |||||||
Db   206 TCAACTTATGCGCGGTTGTACAGAC 231

```

```

RESULT 23
AX884251
LOCUS AX884251 576 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 114 from Patent EP1033401.
ACCESSION AX884251
VERSION AX884251.1 GI:40039225
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 114 06-SEP-2000;
Genset (FR)
FEATURES
    source          Location/Qualifiers
    1..576
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    CDS
    195..>575
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAE988365.1"
    /db_xref="GI:40039226"
    /translations="MRPEDRMFHIRAVILRSLSLATLLSLRGAGALKADHVSTYAAFY
    OTHRPTEGEMFETDEDEMFEYDLDKKEVWHLEFGQSPFPAQGGIANIAIINNIN
    TLIQRNHQAINDPPEVTVPFKP"
    sig_peptide     195..287
    /note="score 10.1 seq SLAFLLSLRGAGA/IK"
    ORIGIN
    Query Match 100.0%; Score 26; DB 6; Length 576;
    Best Local Similarity 100.0%; Pred. No. 1.3;
    Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 TCAACTTATGCCCGGTTGTACAGAC 26
    |||||
    Db 306 TCACTTATGCCCGGTTGTACAGAC 331

RESULT 24
BD023861
LOCUS BD023861 576 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD023861
VERSION BD023861.1 GI:22565084
KEYWORDS JP 2001269182-A/107.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 107 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/107
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PI 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C-2N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//S06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC score 10.1
CC seq SLAFLLSLRGAGA/IK
PH Key Location/Qualifiers

```

```

FEATURES
    source          Location/Qualifiers
    1..576
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    ORIGIN
    Query Match 100.0%; Score 26; DB 6; Length 576;
    Best Local Similarity 100.0%; Pred. No. 1.3;
    Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 TCAACTTATGCCCGGTTGTACAGAC 26
    |||||
    Db 306 TCACTTATGCCCGGTTGTACAGAC 331

RESULT 25
E00485
LOCUS E00485 661 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of a fragment of PSBalpha-318.
ACCESSION E00485
VERSION E00485.1 GI:2168768
KEYWORDS JP 1985226888-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 661)
AUTHORS Edowaado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.
TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE
JOURNAL Patent: JP 1985226888-A 3 12-NOV-1985;
CETUS CORP
COMMENT OS Human (Homo sapiens)
PN CP 1985226888-A/3
PD 12-NOV-1985
PF 20-MAR-1985 JP 1985054705
PF 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI
EDOWAADO RUISU SHIERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI
KOOREI HANAADO REBENSON, HENRII RAPOFORUTO
PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/50,G01N33/532,
PC G01N33/58,
PC (C07D519/00,C07D493:04,C07D495:04),(C07D519/00,C07D493:04, PC
C07D493:10);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=Blood;
CC *source: cell_type=B cell;
CC *source: library=cDNA library;
PH Key Location/Qualifiers
FT misc_feature 1..661
    /note="a fragment derived from pSBalpha-318
    for insertion".
    Location/Qualifiers
    1..661
    /organism="unidentified"
    /mol_type="genomic RNA"
    /db_xref="taxon:32644"
    FEATURES
    source          Location/Qualifiers
    1..661
    /organism="unidentified"
    /mol_type="genomic RNA"
    /db_xref="taxon:32644"
    ORIGIN
    Query Match 100.0%; Score 26; DB 6; Length 661;
    Best Local Similarity 100.0%; Pred. No. 1.2;
    Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 TCAACTTATGCCCGGTTGTACAGAC 26
    |||||
    Db 190 TCACTTATGCCCGGTTGTACAGAC 215

RESULT 26

```



TITLE Isotypic and allotypic variation of human class II histocompatibility antigen alpha-chain genes  
 JOURNAL Nature 308 (5957), 327-333 (1984)  
 MEDLINE 84168117  
 PUBMED 6584734  
 COMMENT On Nov 6, 2003 this sequence version replaced gi:188516.  
 FEATURES  
 source  
 1..1048  
 /location=Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 <1..703  
 /codon\_start=2  
 /product="SB classII histocompatibility antigen alpha-chain"  
 /protein\_id="CAA25143.1"  
 /db\_xref="GI:758100"  
 /db\_xref="COA:P20036"  
 /db\_xref="SWISS-PROT:P20036"  
 /translation="GAGRAIKADHYSTAAVQVTHRPTGEFMEFDEMFVLDLKKETVWHLLEFGAFSFAQGLANLAIILNNLNTLIQRNHTQATNDPPETVTFPPPEVE LGQENTLICHIDKFPFPPVNLVTLNGLVTEGVAESLELPRTDYSFKKHVLTFFVS AEDFYDCRVEHGLDQPLKHWAEQPIQMEPTETVLCALGLVLGLVGVIGVILII KSLRSQHPRAQGL"  
 <1..113  
 /note="signal peptide fragment"  
 14..265  
 /note="alpha 1"  
 266..547  
 /note="alpha 2"  
 548..586  
 /note="(CP) connecting peptide"  
 587..655  
 /note="(TM) transmembrane region"  
 656..700  
 /note="(CY) cytoplasmatic region"  
 1048  
 /note="polyadenylation site"  
 ORIGIN  
 Query Match 100.0%; Score 26; DB 9; Length 1048;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
 Db 32 TCAACTTATGCGCGTTGTACAGAC 57  
 1 TCAACTTATGCGCGTTGTACAGAC 26  
 32 TCAACTTATGCGCGTTGTACAGAC 57  
 RESULT 30  
 I03088  
 LOCUS 1140 bp ss-DNA linear PAT 21-MAY-1993  
 DEFINITION Sequence 7 from Patent US 4582789.  
 ACCESSION I03088  
 VERSION I03088.1 GI:268244  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1140)  
 AUTHORS Sheldon,E.L. III, Levenson,C.H., Mullis,K.B. and Rapoport,H.  
 TITLE Process for labeling nucleic acids using psoralen derivatives  
 JOURNAL Patent: US 4582789-A 7 15-APR-1986;  
 Cetus Corporation, Emeryville, CA  
 FEATURES  
 source  
 1..1140  
 /location=Qualifiers  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 100.0%; Score 26; DB 6; Length 1140;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
 Db 191 TCAACTTATGCGCGTTGTACAGAC 216  
 RESULT 31  
 E00484  
 LOCUS 1201 bp RNA linear PAT 29-SEP-1997  
 DEFINITION DNA sequence of pSBalpha-318.  
 ACCESSION E00484  
 VERSION E00484.1 GI:2168767  
 KEYWORDS JP 1985226888-A/2.  
 SOURCE unidentified  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Edowado,R.S.Z.Z., Kerii,B.M., Koorei,H.R. and Henrii,R.  
 TITLE NUCLEIC ACID LABELLING SUBSTANCE AND USE  
 JOURNAL Patent: JP 1985226888-A 2 12-NOV-1985;  
 CETUS CORP  
 COMMENT OS Human (Homo sapiens)  
 EN JP 1985226888-A/2  
 ED 12-NOV-1985  
 PF 20-MAR-1985 JP 1985054705  
 PR 21-MAR-1984 US 84 591811, 18-DEC-1984 US 84 683263 PI  
 EDOWAADO RUISU SHIERUDON ZA SAADO, KERII BANKUSU MIYURISU, PI  
 KOOREI HAWAADO REBENSON, HENRII RAPOPORJTO  
 PC C07D519/00,C07H21/04,C12N15/00,C12Q1/68,G01N33/532,  
 G01N33/58  
 PC G01N33/58  
 PC (C07D519/00,C07D493/04,C07D495/04),(C07D519/00,C07D493/04, PC  
 C07D493/10);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: tissue\_type=Blood;  
 CC \*source: cell\_type=B cell;  
 CC \*source: library=cDNA library;  
 CC \*source: clone=pSBalpha-318;  
 FH Key Location/Qualifiers  
 FT CDS 1..2201  
 /gene="pSBalpha-318".  
 FT Location/Qualifiers  
 1..1201  
 /organism="unidentified"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:32644"  
 ORIGIN  
 Query Match 100.0%; Score 26; DB 6; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 0.98;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
 Db 191 TCAACTTATGCGCGTTGTACAGAC 216  
 RESULT 32  
 I03006  
 LOCUS 1201 bp ss-DNA linear PAT 21-MAY-1993  
 DEFINITION Sequence 6 from Patent US 4617261.  
 ACCESSION I03006  
 VERSION I03006.1 GI:268462  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Sheldon,E.L. III, Levenson,C.H., Mullis,K.B., Rapoport,H. and  
 Watson,R.M.

TITLE Process for labeling nucleic acids and hybridization probes  
JOURNAL Patent: US 4617261-A 6 14-OCT-1986;  
Cetus Corporation; Emeryville, CA  
FEATURES  
source  
Location/Qualifiers  
1. .1201  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 26; DB 6; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCTTATCGCGGTTTGTACACAC 26  
|||||

Db 191 TCACCTTATCGCGGTTTGTACACAC 216  
|||||

## RESULT 33

LOCUS I03423 1201 bp ss-DNA linear PAT 21-MAY-1993  
DEFINITION Sequence 6 from Patent US 4822731.  
ACCESSION I03423  
VERSION I03423.1 GI:270023  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1201)  
AUTHORS Watson,R.M., Sheldon,E.L. III and Shead,R.M.  
TITLE Process for labeling single-stranded nucleic acids and hybridization probes

JOURNAL Patent: US 4822731-A 6 18-APR-1989;  
Cetus Corporation; Emeryville, CA

## FEATURES

source  
1. .1201  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 26; DB 6; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCTTATCGCGGTTTGTACACAC 26  
|||||

Db 191 TCACCTTATCGCGGTTTGTACACAC 216  
|||||

## RESULT 34

LOCUS BC009956 1201 bp mRNA linear PRI 11-DEC-2003  
DEFINITION Homo sapiens major histocompatibility complex, class II, DP alpha 1, mRNA (cDNA clone MGC:14114 IMAGE:4309471), complete cds.

ACCESSION BC009956

VERSION BC009956.2 GI:39644570

KEYWORDS MGC.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,I., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Pahey,J., Helton,E., Ketterman,M., Madan,A., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)  
12477932

REFERENCE 2 (bases 1 to 1201)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK

## COMMENT

NTH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Dec 9, 2003 this sequence version replaced gi:14602922.

Contact: MGC help desk

Email: [cgabs-x@mail.nih.gov](mailto:cgabs-x@mail.nih.gov)

Tissue Procurement: Louis Staudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,R., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 20 Row: p Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24797073.

## FEATURES

## source

1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:14114 IMAGE:4309471"  
/tissue\_type="Primary B-Cells from Tonsils"  
/clone\_lib="NIH\_MGC\_48"  
/lab\_host="DH10B-R"  
/notes="Vector: pOTB7"

## gene

1. .1201  
/gene="HLA-DPA1"  
/note="Synonyms: HLA-DPIA, HLADP, HLASB"  
/db\_xref="LocusID:3113"  
/db\_xref="MIM:142880"  
32. .814  
/codon\_start=1  
/product="HLA-DPA1 protein"  
/protein\_id="AA00956.1"  
/db\_xref="GI:14602923"  
/db\_xref="LocusID:3113"  
/translation="MRPDRMFHRAVILRALSLAFSLRGAIGAKADSVTAAPV  
TLQRTGEMFEDDEQFVLDKXETVWHLEFGAFSFEAQGLAIANNIN  
CTLRSHHTCAANDPEVTVPKEPVLGPNTLICIDRFPFVNLVTLNGEPTV  
EGVAESLFPRTDYSFKHLYLTFVSAEDVDCRVHGLDQLLKHWAEQFIQMP  
ETTVETVLCALGLVLGVILGVIVGLIKSRSGHDPRAQGL"

## CDS

## misc\_feature

134. .376  
/note="MHC II alpha; Region: Class II histocompatibility antigen, alpha domain"



```

misc_feature      /db_xref="CDD:pfam00993"
374..661
/Note="IG; Region: Immunoglobulin domain constant region
subfamily"
/db_xref="CDD:cd00098"

ORIGIN
Query Match      100.3%; Score 26; DB 9; Length 1201;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
|||||
Db 143 TCAACTTATGCGCGTTGTACAGAC 168

RESULT 35
AX552229          1259 bp DNA linear PAT 27-NOV-2002
LOCUS
DEFINITION
Sequence 138 from Patent WO0162927.
ACCESSION
AX552229
VERSION
AX552229.1 GI:25896647
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Banville,S.C., Greenawalt,L.B., Lincoln,S.E., Stockdreher,T.K.,
Ansley,S., Chang,S.C., Chen,W., D'Sa,S.A., Dam,T.C., Liu,T.F.,
Rosen,B.H., Russo,F.D., Spiro,P.A., Bradley,D.L., Chen,A.,
Cohen,H.J., Daffo,A., Daniels,S.E., Dufour,G.E., Flores,V.,
Fong,W.T., Hodgson,D.M., Jackson,S., Jones,A.L., Panzer,S.,
Roseberry,A.M., Shah,P., Wright,R.J., Yap,P.E., Yu,J.Y.,
Bratcher,S.R., Chalup,M.S., Dahl,C.R. and Hallman,J.L.
Polypeptides and corresponding polynucleotides for diagnostics and
therapeutics
JOURNAL
Patent: WO 0162927-A 138 30-AUG-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..1259
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/Note="Incyte ID No: LI:1169865.1:2000MAY01"

ORIGIN
Query Match      100.0%; Score 26; DB 6; Length 1259;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
|||||
Db 326 TCAACTTATGCGCGTTGTACAGAC 351

RESULT 36
AX780118          1480 bp DNA linear PAT 14-JUL-2003
LOCUS
DEFINITION
Sequence 2275 from Patent WO03039443.
ACCESSION
AX780118
VERSION
AX780118.1 GI:32697112
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
Novel genetic markers for leukemias
JOURNAL
Patent: WO 03039443-A 2275 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE) ;
Ludwig-Maximilian-Universitaet Muenchen (DE) ;
Haferlach, Torsten,
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE) ;

```

```

Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES
source
Location/Qualifiers
1..1480
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 26; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
|||||
Db 400 TCAACTTATGCGCGTTGTACAGAC 425

RESULT 37
AX780119          1480 bp DNA linear PAT 14-JUL-2003
LOCUS
DEFINITION
Sequence 2276 from Patent WO03039443.
ACCESSION
AX780119
VERSION
AX780119.1 GI:32697113
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
Novel genetic markers for leukemias
JOURNAL
Patent: WO 03039443-A 2276 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE) ;
Ludwig-Maximilian-Universitaet Muenchen (DE) ;
Ludwig-Maximilian-Universitaet Muenchen (DE) ;
Haferlach, Torsten,
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE) ;
FEATURES
source
Location/Qualifiers
1..1480
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 26; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
|||||
Db 400 TCAACTTATGCGCGTTGTACAGAC 425

RESULT 38
HUMHDC02          2986 bp DNA linear PRI 07-JAN-1995
LOCUS
DEFINITION
Human MHC class II lymphocyte antigen (DPw4-alpha-1) gene, exons
2-4.
ACCESSION
M23904 J02738 M15446
VERSION
M23904.1 GI:188385
KEYWORDS
cell surface glycoprotein; class II gene; integral membrane
protein; major histocompatibility complex.
SEGMENT
2 of 3
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 2986)
Gustafsson,K., Widmark,E., Jonsson,A.K., Servenius,B., Sachs,D.H.,
Larhammar,D., Rask,L. and Peterson,P.A.
Class II genes of the human major histocompatibility complex.
Evolution of the DP region as deduced from nucleotide sequences of
the four genes

```

J. Biol. Chem. 262 (18), 8778-8786 (1987)  
 MEDLINE 87250502  
 PUBMED 3036829  
 COMMENT Original source text: Human T-cell DNA, clone p412-1 and clone p2703-1.

FEATURES  
 source Location/Qualifiers  
 1..2986  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 join(M23903.1:563..1095,1..2986)  
 /gene="DPw4-alpha-1"  
 order(M23903.1:663..1095,1..769)  
 /gene="DPw4-alpha-1"  
 /number=1  
 770..1015  
 /gene="DPw4-alpha-1"  
 /number=2  
 1016..1355  
 /gene="DPw4-alpha-1"  
 /number=2  
 1356..1637  
 /gene="DPw4-alpha-1"  
 /number=3  
 1638..1851  
 /gene="DPw4-alpha-1"  
 /number=3  
 1852..2018  
 /gene="DPw4-alpha-1"  
 /number=4  
 2019..2986  
 /gene="DPw4-alpha-1"  
 /note="Does not fit consensus"  
 /number=4

ORIGIN About 2.4 kb after segment 1; chromosome 6p21.3.  
 Query Match 100.0%; Score 26; DB 9; Length 2986;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
 Db 781 TCAACTTATGCGCGTTGTACAGAC 806

RESULT 39  
 LOCUS HSHLASBA 14646 bp DNA linear PRI 16-FEB-1995  
 DEFINITION Human HLA-SB(DP) alpha gene.  
 ACCESSION X03100  
 VERSION X03100.1 GI:32243  
 KEYWORDS antigen; cell surface glycoprotein; class II antigen; glycoprotein; inverted repeat; Kpn repetitive sequence; major histocompatibility complex; repetitive sequence.

SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 14646)  
 AUTHORS Lawrence S.K., Das H.K., Pan,J. and Weissman,S.M.  
 TITLE The genomic organization and nucleotide sequence of the HLA-SB(DP) alpha gene  
 JOURNAL Nucleic Acids Res. 13 (20), 7515-7528 (1985)  
 MEDLINE 86041930  
 PUBMED 2997750

COMMENT Data kindly reviewed (05-MAY-1987) by S.K. Lawrence.

FEATURES  
 source Location/Qualifiers  
 1..14646  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 complement(1..247)  
 /gene="HLA-SB beta"

prim\_transcript complement(<1..247)  
 /gene="HLA-SB beta"  
 intron complement(1..78)  
 /gene="HLA-SB beta"  
 /number=1  
 exon complement(79..247)  
 /gene="HLA-SB beta"  
 /number=1  
 CDS complement(<79..178)  
 /gene="HLA-SB beta"  
 /codon\_start=1  
 /protein\_id="CAA26886.1"  
 /db\_xref="GI:32244"  
 /db\_xref="SPTREMBL:O19686"  
 /translation="MMVLQVSAPRTVALTALLMVLLTSVVQGRATP"  
 complement(323..374)  
 /note="beta consensus sequence, put. regulatory region"  
 repeat\_unit 514..526  
 /note="imp. inverted repeat a"  
 repeat\_unit 2189..2200  
 /note="imp. inverted repeat a"  
 misc\_signal 2461..2512  
 /note="alpha consensus sequence; put. regulatory region"  
 gene 2571..11200  
 /gene="HLA-SB"  
 join(2571..2749,6334..6579,6920..7201,7416..7582,10872..11200)  
 /gene="HLA-SB alpha"  
 join(2571..2749,6334..6579,6920..7201,7416..7582,10872..11200)  
 /gene="HLA-SB alpha"  
 prim\_transcript 2571..11200  
 /gene="HLA-SB"  
 exon 2571..2749  
 /gene="HLA-SB"  
 /number=1  
 CDS join(2650..2749,6334..6579,6920..7201,7416..7570)  
 /gene="HLA-SB alpha"  
 /codon\_start=1  
 /product="class II antigen"  
 /protein\_id="CAA26887.1"  
 /db\_xref="GI:673417"  
 /db\_xref="GOA:P20036"  
 /db\_xref="SWISS-PROT:P20036"  
 /translation="MRPDRMFHRAVLRLALSALFLSLRGAGAKADHVSTVAFFVTLQRSHETATNDPPETVFPKPEVFGQAFSGFAQAGLANIAILNNLN EGVASLFLPRDYSPKPHYLPVPSAEDFYDCRVEHGLDQPLLKWEAQEPQMP ETVETVLCALGLVLGVLVGTVLLIKSLRSGHDPRAQGL"  
 2650..2742  
 /gene="HLA-SB"  
 2650..2749  
 /gene="HLA-SB"  
 /note="precursor fragment"  
 2724..2725  
 /gene="HLA-SB"  
 /note="pot. alternate signal sequence splice site"  
 2730..6333  
 /gene="HLA-SB"  
 /number=1  
 4964..4990  
 /gene="HLA-SB"  
 /note="inverted repeat b"  
 5083..5088  
 /gene="HLA-SB"  
 /note="inverted repeat b"  
 6334..6579  
 /gene="HLA-SB"  
 /number=2  
 6580..6919  
 /gene="HLA-SB"  
 /number=2  
 6920..7201  
 /gene="HLA-SB"

```

/gene="HLA-SB"
/number=3
7202..7415
/gene="HLA-SB"
/number=3
7416..7582
/gene="HLA-SB"
/number=4
7583..10871
/gene="HLA-SB"
/number=4
8601..9100
/gene="HLA-SB"
/note="sequence homologous to IgC epsilon genes"
8991..9029
/gene="HLA-SB"
/note="inverted repeat C"
10516..10554
/gene="HLA-SB"
/note="inverted repeat C"
10872..11200
/gene="HLA-SB"
/number=5
12301..12800
/note="Kpn repetitive sequence"

```

```

Query Match      100.0%; Score 26; DB 9; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 26; Conservative C; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TCACTTATGCGCGTTTGTACAGAC 26
Db 6345 TCACTTATGCGCGTTTGTACAGAC 6370

```

```

RESULT 40
BX120009/c
LOCUS BX120009 64380 bp DNA linear PRI 07-AUG-2003
DEFINITION Human DNA sequence from clone DASS-227B13 on chromosome 6, complete
sequence.
ACCESSION BX120009
VERSION BX120009.10 GI:33504469
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Sycamore.N.
JOURNAL Direct Submission
Submitted (07-AUG-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CH10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 7, 2003 this sequence version replaced gi:31335527.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> This sequence was generated from part of bacterial clone contigs constructed by the MHC HaploTYPE Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC> DASS-227B13 is from a DNA-arts SSTS human bac library VECTOR: pBeloeBAC11.

#### FEATURES

```

source
1..64380
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DASS-227B13"
/clone_lib="DNA-arts-BAC.1-SSTO.1"

```

#### ORIGIN

```

Query Match      100.0%; Score 26; DB 9; Length 64380;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TCACTTATGCGCGTTTGTACAGAC 26
Db 2126 TCACTTATGCGCGTTTGTACAGAC 2101

```

```

RESULT 41
AL805913/c

```

```

LOCUS AL805913 106728 bp DNA linear PRI 24-OCT-2002
DEFINITION Human DNA sequence from clone Xxbac-22D21 on chromosome 6, complete
sequence.

```

```

ACCESSION AL805913
VERSION AL805913.4 GI:24395073
KEYWORDS HTG.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tracy, A.
Direct Submission
Submitted (23-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CH10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 25, 2002 this sequence version replaced gi:22204654.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

```

#### COMMENT

```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP

database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs constructed by  
 the MHC Haplotype Consortium and collaborators. Further information  
 can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
 XxBac-22D21 is from a DNA-arts QBL human bac library VECTOR:  
 pBel03AC11.

#### FEATURES

Source  
 Location/Qualifiers  
 1..106728  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXBac-22D21"  
 /clone\_lib="DNA-arts-BAC.1-QBL.1"

#### ORIGIN

Query Match 100.0%; Score 26; DB 9; Length 106728;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACTTATGCGCGTTGTACAGAC 26  
 |||||  
 Db 95130 TCACTTATGCGCGTTGTACAGAC 95105

#### RESULT 42

AL645931/c  
 LOCUS  
 DEFINITION Human DNA sequence from clone XxBac-138A21 on chromosome 6,  
 complete sequence.  
 AL645931  
 AL645931.7 GI:19572887  
 HTG.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

#### REFERENCE

1  
 Almeida, J.  
 Direct Submission  
 Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On Mar 21, 2002 this sequence version replaced gi:19031691.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP  
 database can be found at

[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) XxBac-138A21 is  
 from a CHORI-501 human bac - PGF cell line library VECTOR:  
 pTABEC2.1

This sequence was generated from part of bacterial clone contigs  
 constructed by the MHC Haplotype Consortium and collaborators.  
 Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6/MHC>.

#### FEATURES

Location/Qualifiers  
 1..124899  
 /organism="Homo sapiens"

/mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="XXBac-138A21"  
 /clone\_lib="CHORI-501"

#### ORIGIN

Query Match 100.0%; Score 26; DB 9; Length 124899;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACTTATGCGCGTTGTACAGAC 26  
 |||||  
 Db 77811 TCACTTATGCGCGTTGTACAGAC 77786

#### RESULT 43

AC011086/c  
 LOCUS  
 DEFINITION Homo sapiens chromosome 6 clone RP11-93F3 map 6, WORKING DRAFT  
 SEQUENCE, 12 unordered pieces.  
 AC011086  
 AC011086.5 GI:10047675  
 HTG; HTGS PHASE1; HTGS\_DRAFT.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

#### ACCESSION

VERSION

KEYWORDS

SOURCE

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

#### REFERENCE

1 (bases 1 to 181228)

#### AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

#### TITLE

Homo sapiens chromosome 6, clone RP11-93F3

#### JOURNAL

Unpublished

#### REFERENCE

2 (bases 1 to 181228)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouckgalter, B.,  
 Brown, A., Castle, A., Collangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

#### Direct Submission

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

#### REFERENCE

3 (bases 1 to 181228)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Beka, F., Boguslavsky, L.,  
 Bouckgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,  
 Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,  
 Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,  
 Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,  
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Lakocque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,  
 Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,  
 McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,  
 O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,  
 Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,  
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
 Wilson, S., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,  
 Zimmer, A. and Zody, M.

#### Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome

## COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 10, 2000 this sequence version replaced gi:7717099.  
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3322

Center clone name: 93\_F\_3

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 3.960731

Consensus quality: 166135 bases at least Q40

Consensus quality: 173448 bases at least Q30

Consensus quality: 176936 bases at least Q20

Insert size: 174000; agarose-fp

Insert size: 180128; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 27639: contig of 27639 bp in length  
\* 27640 27739: gap of 100 bp  
\* 27740 28970: contig of 1331 bp in length  
\* 28971 29070: gap of 100 bp  
\* 29071 82294: contig of 53224 bp in length  
\* 82295 82394: gap of 100 bp  
\* 82395 87214: contig of 4820 bp in length  
\* 87215 87314: gap of 100 bp  
\* 87315 93029: contig of 5715 bp in length  
\* 93030 93129: gap of 100 bp  
\* 93130 101779: contig of 8850 bp in length  
\* 101780 101879: gap of 100 bp  
\* 101880 108114: contig of 6235 bp in length  
\* 108115 108214: gap of 100 bp  
\* 108215 116143: contig of 7929 bp in length  
\* 116144 116243: gap of 100 bp  
\* 116244 133838: contig of 17595 bp in length  
\* 133839 133938: gap of 100 bp  
\* 133939 153417: contig of 19479 bp in length  
\* 153418 153517: gap of 100 bp  
\* 153518 178444: contig of 24927 bp in length  
\* 178445 178544: gap of 100 bp  
\* 178545 182228: contig of 2684 bp in length.

FEATURES  
source

Location/Qualifiers  
1..181228  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6"  
/clone="RP11-93F"

/clone\_lib="RP11-11 Human Male BAC"

1..27639

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

27740..28970

/note="assembly\_fragment"

29071..82294

/note="assembly\_fragment"

misc\_feature

82395..87214

/note="assembly\_fragment"

misc\_feature

87315..93029

/note="assembly\_fragment"

misc\_feature

93130..101779

/note="assembly\_fragment"

misc\_feature

101880..108114

/note="assembly\_fragment"

misc\_feature

108215..116143

/note="assembly\_fragment"

misc\_feature

116244..133838

/note="assembly\_fragment"

misc\_feature

133939..153417

/note="assembly\_fragment"

misc\_feature

153518..178444

/note="assembly\_fragment"

misc\_feature

178445..181228

/note="assembly\_fragment"

misc\_feature

clone\_end:T7

vector\_side:right

## ORIGIN

Query Match 100.0%; Score 26; DB 2; Length 181228;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGGTTGTACAGAC 26

|||||

Db 64049 TCAACTTATGCGCGGTTGTACAGAC 64024

## RESULT 44

AL662824/c

LOCUS AL662824

DEFINITION Human DNA sequence from clone Xxbac-22116 on chromosome 6, complete sequence.

ACCESSION AL662824

VERSION AL662824.9 GI:20068657

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM

Sukaryota; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Almeida,J.

Direct Submission

Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Apr 7, 2002 this sequence version replaced gi:20067828.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em:, EMBL; Sw:,

SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep Xxbac-22116 is

from a CHORI-502 human bac - COX cell line library VECTOR:

PTARBAC2.1

This sequence was generated from part of bacterial clone contigs

constructed by the MHC RapiType Consortium and collaborators.

Further information can be found at

misc\_feature

misc\_feature

misc\_feature

FEATURES	Direct Submission	
	Submitted (15-SEP-1998) Tissue Typing, Royal Melbourne Hospital, Grattan Street, Parkville, VIC 3050, Australia	
	Location/Qualifiers	
source	1..265	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
gene	/isolate="904"	
	/db_xref="taxon:9606"	
	/chromosome="6"	
mRNA	<1..>265	
	/gene="HLA-DPA1"	
	<1..>265	
CDS	/gene="HLA-DPA1"	
	/codon_start=3	
	/product="MHC class II antigen"	
exon	/protein_id="AAF00051.1"	
	/db_xref="GI:6002597"	
	/translation="VSTYAMFYQTHPTGGEVFPEDEQFYVDLKKETVWHLPEFG RAFSFEAQGLANIALNNLNLLIQRSHNTQAANGTAYLCFL"	
ORIGIN	1..265	
	/gene="HLA-DPA1"	
	/number=2	
Query Match	87.7%; Score 22.8; DB 9; Length 265;	
	Best Local Similarity 92.3%; Pred. No. 42;	
	Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Db	1 TCAACTTATGCGCGTTGTACAGAC 26	
	6 TCAACTTATGCGCGTTGTACAGAC 31	
RESULT 47	AF013767 326 bp DNA linear PRI 16-OCT-1998	
	AF013767 Homo sapiens MHC class II antigen HLA-DPA1 gene (HLA-DPA1*0302	
	allele); partial cds.	
ACCESSION	AF013767 U94838	
	AF013767.1 GI:3660651	
KEYWORDS	Homo sapiens (human)	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 326)	
	Steiner,L.L., Cavalli,A., Zimmerman,P.A., Boatin,B.A.,	
AUTHORS	Titanji,V.P., Bradley,J.E., Lucius,R., Nutman,T.B. and	
	Begovich,A.B.	
	Three new DP alleles identified in sub-Saharan Africa: DPB1*7401,	
TITLE	DPB1*02013, and DPB1*0302	
	Tissue Antigens 51 (6), 653-657 (1998)	
	98357732	
JOURNAL	MEDLINE	
	PUBMED	
	9694359	
REFERENCE	2 (bases 1 to 326)	
	Steiner,L., Begovich,A. and Zimmerman,P.	
	Direct Submission	
TITLE	Submitted (14-JUL-1997) Human Genetics, Roche Molecular Systems,	
	1145 Atlantic Ave., Alameda, CA 94501, USA	
	On Oct 16, 1998 this sequence version replaced gi:2865247.	
COMMENT	location/Qualifiers	
	1..326	
	/organism="Homo sapiens"	
FEATURES	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="6"	
source	<1..>326	
	/gene="HLA-DPA1"	
	/allele="HLA-DPA1*0302"	
gene		
FEATURES	http://www.sanger.ac.uk/HGP/Chr6/MHC.	
	Location/Qualifiers	
	1..187964	
source	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
gene	/chromosome="6"	
	/clone_lib="XXbac-22116"	
	/clone_lib="CHORI-502"	
ORIGIN	100.0%; Score 26; DB 9; Length 187964;	
	Query Match	
	Best Local Similarity 100.0%; Pred. No. 317;	
Matches	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 TCAACTTATGCGCGTTGTACAGAC 26	
	51840 TCAACTTATGCGCGTTGTACAGAC 51815	
RESULT 45	HUMHCH1A	
	LOCUS	
	DEFINITION	
ACCESSION	L11641	
	VERSION	
	L11641.1 GI:187872	
KEYWORDS	alpha-subunit; class II gene; major histocompatibility complex.	
	Homo sapiens (human)	
	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 214)	
REFERENCE	Guethlein,L.A., Bias,W.B. and Schmeckpeper,B.J.	
	New DP sequences: three DPA1 and one DPB1	
	Tissue Antigens 41 (5), 269-272 (1993)	
TITLE	94054425	
	MEDLINE	
	PUBMED	
COMMENT	8236240	
	Original	
	source text: Homo sapiens DNA.	
FEATURES	1..214	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
source	/db_xref="taxon:9606"	
	/cell_line="K117 (1099024)"	
ORIGIN	87.7%; Score 22.8; DB 9; Length 214;	
	Query Match	
	Best Local Similarity 92.3%; Pred. No. 45;	
Matches	24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
	24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
	24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Db	1 TCAACTTATGCGCGTTGTACAGAC 26	
	11 TCAACTTATGCGCGTTGTACAGAC 36	
RESULT 46	AF092049	
	LOCUS	
	DEFINITION	
ACCESSION	AF092049	
	VERSION	
	AF092049.1 GI:6002596	
KEYWORDS	Homo sapiens (human)	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 265)	
	Varney,M.D., Gavrilidis,A. and Abbott,W.	
AUTHORS	DPB1 Polymorphism in Polynesians	
	Unpublished	
	JOURNAL	
TITLE	2 (bases 1 to 265)	
	Varney,M.D., Gavrilidis,A. and Abbott,W.	
	REFERENCE	

```

mRNA
<1..>326
/gene="HLA-DPA1"
/product="MHC class II antigen"
CDS
<1..>326
/codon_start=1
/gene="HLA-DPA1"
/product="MHC class II antigen"
/protein_id="AAC54233.1"
/db_xref="GI:3757788"
/translation="TFGRVFLYFLYADHVSYAMFVOTHTPTGTFEPEDEMFYV
DLDKKETVHLERFQAFPEAQGLANIALNNLNLTLIORSNHTQATNGTPYICLF
LCSPSTG"
ORIGIN
Query Match 87.7%; Score 22.8; DB 9; Length 326;
Best Local Similarity 92.3%; Pred. No. 39;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTTGTACAGAC 26
|||||
DB 49 TCAACTTATGCAATGTTGTACAGAC 74
|||||

RESULT 48
AX344560/c
LOCUS
DEFINITION Sequence 11 from Patent WO200932.
ACCESSION AX344560
VERSION AX344560.1 GI:18492446
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Olek, A., Piepenbrock, C. and Berlin, K.
AUTHORS
TITLE Diagnosis of known genetic parameters within the mhc
JOURNAL Patent: WO 0200932-A 11 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
1..349980
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3,673,778 <223>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3,673,778 <223>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"
ORIGIN
Query Match 81.5%; Score 21.2; DB 6; Length 349980;
Best Local Similarity 88.5%; Pred. No. 17;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTTGTACAGAC 26
|||||
DB 298050 TCAACTTATACCGCGTTTATACAAAC 298025
|||||

RESULT 49
HUMHCIIAY

```

```

LOCUS
DEFINITION Human MHC class II (HLA-DPA1) gene.
ACCESSION L11642
VERSION L11642.1 GI:187873
KEYWORDS alpha-subunit; class II gene; major histocompatibility complex.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 220)
AUTHORS Guethlein, L.A., Bias, W.B. and Schmeckepeper, B.J.
TITLE New DP sequences: three DP1 and one DP1
JOURNAL Tissue Antigens 41 (5), 269-272 (1993)
MEDLINE 94054425
PUBMED 8236240
COMMENT Original source text: Homo sapiens DNA.
FEATURES
Location/Qualifiers
1..220
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/cell_line="CB6B (10M9060)"
ORIGIN
Query Match 80.0%; Score 20.8; DB 9; Length 220;
Best Local Similarity 91.7%; Pred. No. 3.3e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACTTATGCGCGTTTGTACAGAC 26
|||||
DB 1 AACTTATGCCATGTTGTACAGAC 24
|||||

RESULT 50
HSSBA2P
LOCUS
DEFINITION Human HLA-SB alpha 2 pseudogene (DP-alpha).
ACCESSION X01634 X01407
VERSION X01634.1 GI:36411
KEYWORDS direct repeat; histocompatibility antigen; pseudogene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1834)
AUTHORS Serenius, B., Gustafsson, K., Widmark, E., Emmoth, E., Andersson, G., Larhammar, D., Rask, L. and Peterson, P.A.
TITLE Molecular map of the human HLA-SB (HLA-DP) region and sequence of an SB alpha (DP alpha) pseudogene
JOURNAL EMBO J. 3 (13), 3209-3214 (1984)
MEDLINE 85126928
PUBMED 6597088
COMMENT Data kindly reviewed (12-FEB-1986) by B. Serenius Splice junctions refer to expressed SB alpha gene (Auffray et al 1984. Nature 308, 327-333).
pos. 398 to 399: one base deleted in pseudogene;
pos. 423 to 425: mutated donor splice junction in pseudogene; pos.
1142 to 1144: compensatory mutation to Cys in pseudogene; pos.
1148 to 1150: mutation Cys to Leu in pseudogene; pos. 1533 to
1534: one base deleted in pseudogene; pos. 1548 to 1550: mutation
causing premature stop codon in pseudogene.
FEATURES
Location/Qualifiers
1..1834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..177
/note="intron 1 fragment"
join(178..422,915..1186,1412..1562)
/pseudo
/codon_start=3
/product="HLA-SB alpha 2 pseudogene (DP-alpha)"
178..422
intron
CDS
exon

```

```

/note="exon 1 coding region (178 is 2nd base in codon)
(422 is 1st base in codon)"
/number=1
repeat_region 337..447
/note="direct repeat"
intron 423..914
/note="intron II"
/number=1
repeat_region 448..558
/note="direct repeat"
exon 915..1196
/note="exon 2 coding region (915 is 2nd base in codon)
(1196 is 1st base in codon)"
/number=2
intron 1197..1411
/note="intron III"
/number=2
exon 1412..1574
/note="exon 3 coding region (1412 is 2nd base in codon)"
/number=3
intron 1575..>1834
/note="intron IV fragment"
/number=3

ORIGIN
Query Match 75.4%; Score 19.6; DB 9; Length 1834;
Best local similarity 84.6%; Pred.No. 5.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAACTTATCGCGGTTGTACAGAC 26
||||| ||||| ||||| |||||
Db 189 TCAACATATCGGAGTTGTACAGAC 214

RESULT 51
HSD0826Q1
LOCUS
DEFINITION
HSD0826Q1 22847 bp DNA linear PRI 19-SEP-2001
Human DNA sequence from clone XX-CD0826Q1 on chromosome 6. Contains
a pseudogene similar to part of collagen type XI alpha 2 (COL11A2),
the major histocompatibility complex class II DP alpha 2 pseudogene
HLA-DPA2, ESTs, STSs and GSSs, complete sequence.
AL049813
AL049813.1 GI:4775644
FTG; COL11A2; collagen; HLA-DPA2; MHC.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Williams,S.
Direct Submission
Submitted (19-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human

```

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group in collaboration with Jethro Herberg and John Trowsdale. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> IMPORTANT: This sequence is not the entire insert of clone XX-CD0826Q1. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone XX-LC11 is at 22748 in this sequence.  
The true right end of clone RP5-1033B10 is at 100 in this sequence.

```

FEATURES
    source
        1..22847
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /clone="XX-CD0826Q1"
    repeat_region
        1..579
            /note="L1PB1 repeat: matches 5524..6145 of consensus"
    repeat_region
        693..837
            /note="L2 repeat: matches 1290..1435 of consensus"
    repeat_region
        1200..1510
            /note="L1MA8 repeat: matches 5999..6288 of consensus"
    repeat_region
        1761..2284
            /note="L2 repeat: matches 2120..2680 of consensus"
    misc_feature
        complement(2291..2721)
            /note="match: GSS: Em:AQ694429"
    gene
        join(3549..3685,3801..3970,4459..4610,6424..6468,
        6594..6653,6760..6797)
            /gene="cd0826Q1.1"
    /pseudo
    join(3549..3685,3801..3970,4459..4610,6424..6468,
    6594..6653,6760..6797)
        /gene="cd0826Q1.1"
    /note="cd0826Q1.1 (pseudogene similar to part of collagen
    type XI alpha 2 (COL11A2))"
    match: proteins: Tr:Q99866 Sw:P13942 Tr:Q61432 Sw:Q64739
    Sw:P2107 Tr:Q90589 Tr:O88207 Sw:Q61245 Sw:P20508
    Tr:Q62750 Tr:O57580 Tr:Q921W0 Sw:Q62280"
    /pseudo
    /codon_start=1
    /evidence=not_experimental
    repeat_region
        4186..4351
            /note="MIR repeat: matches 65..261 of consensus"
    repeat_region
        6351..7315
            /note="L1PB2 repeat: matches 5776..6154 of consensus"
    misc_feature
        7257..7409
            /note="match: GSS: Em:AQ688440"
    repeat_region
        7399..8718
            /note="L1PB2 repeat: matches 4435..5791 of consensus"
    repeat_region
        8719..9679
            /note="MER11C repeat: matches 1..1067 of consensus"
    repeat_region
        9686..12936
            /note="L1PB2 repeat: matches 1099..4439 of consensus"
    repeat_region
        12935..13961
            /note="L1PBa repeat: matches -1207..-194 of consensus"
    gene
        join(13269..13323,17426..17782,18163..18444,18660..18810)
            /gene="HLA-DPA2"
    /pseudo
    join(13269..13323,17426..17782,18163..18444,18660..18810)
        /gene="HLA-DPA2"
    /note="cd0826Q1.2 (major histocompatibility complex, class
    II, DP alpha 2 (pseudogene))"
    match: proteins: Tr:Q50181 Sw:P20036 Sw:P20755 Tr:Q31287
    Tr:Q30437 Tr:O19810 Tr:O19434 Tr:Q30761 Tr:Q31064
    Tr:Q31065 Tr:Q30302 Sw:P01906 Tr:Q07455
    match: genomic DNA: Em:X01634"
    /pseudo
    /codon_start=1
    /evidence=not_experimental
    repeat_region
        13962..14259
            /note="Alu repeat: matches 2..298 of consensus"
    repeat_region
        14260..14500
            /note="L1PBa repeat: matches -1449..-1207 of consensus"

```



```

repeat_region 14768..14894
/note="FAM.C repeat: matches 1..127 of consensus"
repeat_region 1580..15845
/note="L2 repeat: matches 1995..2283 of consensus"
repeat_region 17042..17103
/note="MIR repeat: matches 79..144 of consensus"
repeat_region 17140..17171
/note="L2 repeat: matches 2703..2734 of consensus"
repeat_region 17585..17806
/note="2 copies 111 mer 100% conserved"
repeat_region 18959..19182
/note="MIR repeat: matches 11..262 of consensus"
repeat_region 19187..19339
/note="L1P repeat: matches 4059..4223 of consensus"
misc_feature 19433..19937
/note="match: GSS: Em:AQ814971"
misc_feature 19568..19896
/note="match: SRS: Em:G55743
match: GSS: Em:AQ323676"
repeat_region 19597..19954
/note="L1M4 repeat: matches 3345..3710 of consensus"
misc_feature 20063..20156
/note="match: GSS: Em:AQ297034"
repeat_region 20704..20754
/note="47 copies 2 mer tt 62% conserved"
repeat_region 20704..20754
/note="L1ME3 repeat: matches 6030..6080 of consensus"
repeat_region 20808..20873
/note="L2 repeat: matches 2675..2739 of consensus"
repeat_region 20933..21171
/note="MT11 repeat: matches 72..323 of consensus"

ORIGIN
Query Match 75.4%; Score 19.6; DB 9; Length 22847;
Best Local Similarity 84.6%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAACTTATCGCGTGTGTACAGAC 26
||||| ||||| ||||| ||||| |||||
Db 17437 TCAACATATCGGAGTTGTGCAGAC 17462

RESULT 52
AL845446/c
LOCUS
DEFINITION
Human DNA sequence from clone XXbac-79P13 on chromosome 6, complete sequence.
ACCESSION AL845446
VERSION AL845446.7 GI:24474482
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 93842)
Johnson,C.
Direct Submission
Submitted (31-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquery@sanger.ac.uk
On Nov 1, 2002 this sequence version replaced GI:24430255.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hamquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

```

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep
This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC
XXbac-79P13 is from a DNA-arts QBL human bac library VECTOR: pBelobAC11.

FEATURES
            source
            1..93842
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosomes="6"
            /clone_lib="DNA-arts-BAC.1-QBL.1"

ORIGIN
Query Match 75.4%; Score 19.6; DB 9; Length 93842;
Best Local Similarity 84.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAACTTATCGCGTGTGTACAGAC 26
||||| ||||| ||||| ||||| |||||
Db 13655 TCAACATATCGGAGTTGTGCAGAC 13630

RESULT 53
AL645940/c
LOCUS
DEFINITION
Human DNA sequence from clone XXbac-157A10 on chromosome 6, complete sequence.
ACCESSION AL645940
VERSION AL645940.4 GI:18564736
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Leongamornlert,D.
Direct Submission
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquery@sanger.ac.uk
On Feb 6, 2002 this sequence version replaced GI:18070925.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

```

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep/xxbac-157A10 is from a CHORI-501 human bac - PGF cell line library VECTOR: PTABAC2.1

This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/MHC.

## FEATURES

```

source
1..158033
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="XXbac-157A10"
/clone_lib="CHORI-501"

```

## ORIGIN

```

Query Match      75.4%; Score 19.6; DE 9; Length 158033;
Best Local Similarity 84.6%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 TCAACTTATGCGCGTTTCTACAGAC 26

Db 16368 TCACGTATGCAGAGTTGTGCAGAC 16343

## RESULT 54

HS1033B10

LOCUS

DEFINITION

Human DNA sequence from clone RPS-1033B10 on chromosome 6p21.2-21.31. Contains the C6ORF11 gene for chromosome 6 open reading frame 11 (BING4), the B3GALT4 gene for UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase polypeptide 4, the RPS18 gene for 40S ribosomal protein S18, the SACM2L gene for SAC2 (suppressor of actin mutation 2, yeast, homolog)-like (ARE31), an HIV TAT specific factor 1 (HRAUSF1) pseudogene, a zinc finger pseudogene, the RING1 gene for ring finger protein 1 (RNFI), the FABG1 gene for Fabg (beta-ketoadyl-lacyl-carrier-protein) reductase, E coli like protein, the HKE4 gene for HLA class II region protein K24, the RXRB gene for retinoid X receptor beta, the COL11A2 gene for collagen type XI alpha 2, an HLA class II histocompatibility antigen D or S beta pseudogene and three novel genes. Contains seven CpG islands, ESTs, STSs, and GSSs, complete sequence.

## ACCESSION

AL031228

AL031228.1

BTG; ARE1; B3GALT4; BING4; C6ORF11; COL11A2; collagen; CpG island;

FABG1; HKE4; HRAUSF1; retinoid X receptor beta; ribosomal protein;

ring finger; RING1; RING2; RNFI; RPS18; RXRB; SACM2L; zinc finger.

Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 175737)

Tubby2.

Direct Submission

Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

Requests: clonerequest@sanger.ac.uk

On Sep 24, 1998 this sequence version replaced gi.3550015.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phased quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6

RPS-1033B10 is from the library RPS-1033B10 constructed by the group of

Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pCYPAC2

This sequence is the entire insert of clone RPS-1033B10.

## FEATURES

## source

```

1..175737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPS1033B10"
/db_xref="taxon:9606"
/chromosome="6"
/map="p21.2-21.31"
/clone="RPS-1033B10"
/clone_lib="RPS-1033B10"

```

repeat\_region

```
1..61
/note="Alu repeat: matches 2..62 of consensus"
```

repeat\_region

```
63..368
/note="AluSx repeat: matches 1..299 of consensus"
```

repeat\_region

```
374..403
/note="Cheshire repeat: matches 2264..2292 of consensus"
```

repeat\_region

```
404..708
/note="AluSx repeat: matches 1..305 of consensus"
```

repeat\_region

```
719..1027
/note="AluSx repeat: matches 1..303 of consensus"
```

repeat\_region

```
1028..1220
/note="Cheshire repeat: matches 2055..2256 of consensus"
```

repeat\_region

```
1227..1583
/note="Cheshire repeat: matches 1..376 of consensus"
```

repeat\_region

```
1593..1671
/note="MIR repeat: matches 162..241 of consensus"
```

repeat\_region

```
1962..2265
/note="AluSx repeat: matches 1..296 of consensus"
```

repeat\_region

```
2305..2609
/note="AluSx repeat: matches 1..307 of consensus"
```

repeat\_region

```
2749..2926
/note="AluSx repeat: matches 1..165 of consensus"
```

repeat\_region

```
3009..3297
/note="AluSx repeat: matches 1..290 of consensus"
```

repeat\_region

```
4125..4385
/note="AluSx repeat: matches 20..294 of consensus"
```

repeat\_region

```
5020..5113
/note="MIR repeat: matches 110..214 of consensus"
```

gene

```
5240..7122
/gene="C6ORF11"
```

mRNA

```
join(<5240..5553,5705..5799,6368..6463,6617..>7122)
```

```
/gene="C6ORF11"
```

```
/product="dJ1033B10.2.2 (chromosome 6 open reading frame
```

```
11 (BING4), isoform 2)"
```

```
/note="Continued from c1CF0811.6 in Em:297184
```

```
match: ESTs: Em:AW193506 Em:BB617532"
```

```
/evidence=not experimental
```

```
join(<5240..5553,5705..5799,6368..6463,6617..>7122,
```

```
6853..7122)
```

```
/gene="C6ORF11"
```

```
/product="dJ1033B10.2.1 (chromosome 6 open reading frame
```

```
11 (BING4), isoform 1)"
```

```
/note="Continued from c1CF0811.6 in Em:297184
```

```
match: CDNAS: Em:AK009209
```

```
match: ESTs: Em:AL204504 Em:BF338950 Em:CNSL1D01
```

```
Em:AW192492 Em:BG334317 Em:BG476159 Em:CNSL1FLIA"
```

```
/evidence=not experimental
```

```
join(<5240..5553,5705..5799,6368..6463,6617..>7122)
```

```
/gene="C6ORF11"
```

```
/codon_start=2
```

CDS



```
/evidence=not_experimental
/product="d1033B10.2.2 (chromosome 6 open reading frame
11 (BING4), isoform 2)"
/protein_id="CAC88173.2"
/db_xref="GI:15799226"
/db_xref="SPTRMBL:O96KL2"
/translation="YNATSGLBHQLKIFDLRGTQPLSTRTIPHGAGHIAFQSGRLV
AGMDVNIWAGQKASPSLPQYLTHRLSGPVHGLQCPFDVLGVGHTGGITSMV
VFGAGPNFDGLESNPSYRKQKQWEVKALLEKPAELICLDPRALAEVDVISEQG
KKEQIERLGYDPAKAPQPKQKGRSTASIVKEKRVMDDEERHVNHEHGLPQ
AAPSCLCLCPQAPASVSVSPSPVPRTRSGRAFSSIIIRRRSPRPGPHLWID
LCAPSDSRVAVQSLPKICTREMSVFNKEVAVPLPQLGVDSCLLGGWVIGKESDF
LD"
join(<5240..5553,5705..5799,6368..6463,6617..6730,
6853..6951)
/gene="C6ORF11"
/note="match: proteins: Tr:O:5213"
/codon_start=2
/evidence=not_experimental
/product="d1033B10.2.1 (chromosome 6 open reading frame
11 (BING4), isoform 1)"
/protein_id="CAC38438.2"
/db_xref="GI:15799225"
/db_xref="SPTRMBL:O96KL3"
/translation="YNATSGLBHQLKIFDLRGTQPLSTRTIPHGAGHIAFQSGRLV
AGMDVNIWAGQKASPSLPQYLTHRLSGPVHGLQCPFDVLGVGHTGGITSMV
VFGAGPNFDGLESNPSYRKQKQWEVKALLEKPAELICLDPRALAEVDVISEQG
KKEQIERLGYDPAKAPQPKQKGRSTASIVKEKRVMDDEERHVNHEHGLPQ
KKAQAKPTGARSALDRVR"
5838..6062
/note="9 copies 25 mer 88% conserved"
complement(6837..7118)
/note="match: STS: Em:G23698"
7094..7099
/gene="C6ORF11"
7111
/gene="C6ORF11"
7119
/gene="C6ORF11"
7122
/gene="C6ORF11"
complement(7409..9031)
/gene="B3GALT4"
complement(7409..9031)
/gene="B3GALT4"
/product="d1033B10.3 (UDP-Gal:betaGalNAc beta
1,3-galactosyltransferase, polypeptide 4)"
/note="match: cDNAs: Em:Y15061 Em:AB003478 Em:AB026730
match: BSTS: Em:AA365900 Em:H20531 Em:R40008 Em:AA987754
Em:AA662961 Em:AA878755 Em:H20623 Em:R13948"
/evidence=not_experimental
complement(7409)
/gene="B3GALT4"
complement(7416..7421)
/gene="B3GALT4"
7419..7588
/note="match: STS: Em:G43302"
complement(7671..8807)
/gene="B3GALT4"
/note="match: proteins: Tr:O54904 Tr:O43825 Tr:O54905
Tr:O54906 Tr:O24157"
/codon_start=1
/evidence=not_experimental
/product="d1033B10.3 (UDP-Gal:betaGalNAc beta
1,3-galactosyltransferase, polypeptide 4)"
/protein_id="CAA20230.1"
/db_xref="GI:3820979"
/db_xref="GOA:O96024"
/db_xref="SWISS-PROT:O96024"
/translation="NQLRFLRLALALLVIVMTLFGPSGLGEELSLSLASLLPAP
ASPGPPLALPILPNQACSGPGAPFLLIIVCTAPENLORNARIRASWGGLREARG
LRVOTFLFLGPNAPHPYWGSGSLASAAQGDILCAAFQDSYVNLTKTISGLNW
AEKHCPMARYVLTKDDDDYVNVNPELVSELVLRGGWQWERSRTPQREAEQGGQVLH
```

misc\_feature

gene

mRNA

Query Match 75.4%; Score 19.6; DB 9; Length 175737;  
Best Local Similarity 84.6%; Pred. No. 1.1e+02;  
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCACACTTATGCGCGCTTTGTACAGAC 26  
Db 154897 TCACGTATGACAGTTTGTGACGAC 154922

RESULT 55  
AX344564

LOCUS AX344564 349980 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 15 from Patent WO2000332.  
ACCESSION AX344564  
VERSION AX344564.1 GI:18492450  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.

REFERENCE

1 Olex,A., Piepenbrock,C. and Berlin,X.  
Diagnosis of known genetic parameters within the mhc  
Patent: WO 0200932-A 15 03-JAN-2002;  
Epigenomics AG (DE)

FEATURES

Location/Qualifiers  
1..349980  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.673778 <23>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 1.249.980-seq 03 600.001 949.980-seq 04 900.001 1.849.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 2.449.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <23>-Original length of seq 2: 3.673778 <22>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 1.249.980-seq 16 600.001 949.980-seq 17 900.001 1.849.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"

ORIGIN

Query Match 75.4%; Score 19.6; DB 6; Length 349980;  
Best Local Similarity 84.6%; Pred. No. 87;  
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCACACTTATGCGCGCTTTGTACAGAC 26  
Db 75729 TTAATTATGTCGCGTTTGTATAGAC 75754

RESULT 56

AX391467  
LOCUS AX391467 2071 bp mRNA linear VRT 07-OCT-2003  
DEFINITION Danio rerio TGF-beta-induced factor (TGIF) mRNA, complete cds.  
ACCESSION AX391467  
VERSION AX391467.1 GI:37362301



RESULT 58  
 LOCUS AL845258/c 164325 bp DNA linear ROD 05-OCT-2002  
 DEFINITION Mouse DNA sequence from clone RP23-315H12 on chromosome 2, complete  
 sequence.  
 ACCESSION AL845258  
 VERSION AL845258.4 GI:23559344  
 KEYWORDS HTG  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 164325)  
 TITLE Almeida, J.  
 JOURNAL Direct Submission  
 Submitted (05-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Oct 7, 2002 this sequence version replaced gi:23337803.  
 COMMENT ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) RP23-315H12 is  
 from the RP21-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACE3.6.  
 Location/Qualifiers  
 1..164325  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="RP23-315H12"  
 /clone\_lib="RP21-23"

ORIGIN  
 Query Match 71.5%; Score 18.6; DB 10; Length 164325;  
 Best Local Similarity 84.0%; Pred No. 3.1e-02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 CAACCTATGCCGCGTTGTACAGAC 26  
 |||||  
 Db 102309 CAACCTATGCCGCTTTGTACAGAC 102285  
 |||||

RESULT 59  
 AC119677  
 LOCUS Homo sapiens chromosome 1 clone RP5-1066H13, complete sequence.  
 DEFINITION AC119677  
 AC119677 AL15843  
 VERSION AC119677.2 GI:22507128

KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 170397)  
 AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
 Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and  
 Haugen, E.D.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 170397)  
 AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-APR-2002) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 REFERENCE 3 (bases 1 to 170397)  
 AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
 Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and  
 Haugen, E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-AUG-2002) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 COMMENT On Aug 27, 2002 this sequence version replaced gi:20340503.  
 ----- Genome Center  
 Center: University of Washington Genome Center  
 Center Code: UWGC  
 Web site: <http://www.genome.washington.edu>  
 Contact: [uwgctgs@u.washington.edu](mailto:uwgctgs@u.washington.edu)  
 Drafting Center: SC  
 ----- Project Information  
 Center project name: Chr-1  
 Center clone name: RP5-1066H13 (sc0855)  
 ----- Summary Statistics  
 Sequencing vector: plasmid; 54% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 170352 bases at least Q40  
 Consensus quality: 170395 bases at least Q30  
 Consensus quality: 170397 bases at least Q20  
 Insert size: 170397; sum-of-contigs  
 Quality coverage: 9.0x in Q20 bases; sum-of-contigs

Overlapping Sequences:  
 5': RP11-72N2 (UWGC:sc0108) AL354914  
 3': RP11-348A7 (UWGC:sc0674) AL391730  
 -----  
 Sequence Quality Assessment:  
 This entry has been annotated with sequence quality  
 estimates computed by the Phrap assembly program.  
 All manually edited bases have been reduced to quality zero.  
 Quality levels above 40 are expected to have less than  
 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the  
 Genbank flat file format but are available as part  
 of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
 all regions were either double-stranded or sequenced with an  
 alternate chemistry or covered by high quality data (i.e., Phred  
 quality >= 30); an attempt was made to resolve all sequencing  
 problems, such as compressions and repeats; all regions were  
 covered by at least one plasmid subclone or more than one M13  
 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:  
 This sequence has been validated by Multiple Complete Digest  
 fingerprinting. Comparison of the experimentally derived digest  
 fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and  
 vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-800 bp)

are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgII		EcoRI		NsiI	
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
9834	9501	3142	3147	11592	11142
5671	5593	2184	2111	3357	3377
11496	11304	8065	8028	266	<800
6899	6863	486	<800	1629	1629
1829	1850	106	<800	4730	4732
3117	3149	6064	5990	378	<800
1479	1447	978	1031	904	898
7344	7327	6375	6362	180	<800
9694	9501	9624	9528	1332	1311
1166	1116	2971	3008	5789	5832
868	873	697	<800	8560	8619
1036	1116	786	783	5458	5501
3456	3474	4277	4204	6763	6822
3835	3852	4159	4204	6062	6181
75	<800	4281	4204	6897	6922
3299	3335	8055	8028	649	<800
363	<800	1630	1579	6184	6181
2967	2984	6539	6591	1973	1990
2407	2441	1043	1031	8153	8231
12198	12021	1700	1656	2218	2218
8448	8585	334	<800	1727	1629
9192	9501	1849	1812	21292	21661
1905	1850	1510	1449	1005	1005
4161	4121	3399	3370	7052	7067
375	<800	5301	5295	1674	1629
1002	1010	2101	2111	8048	8231
3609	3644	6989	7063	3417	3496
4138	4121	3928	3851	17236	17278
525	<800	1999	1955	2766	2823
1248	1244	3185	3147	15442	15198
2242	2308	10523	10303	1619	1629
3864	3852	19148	19438	11558	11142

-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----		-----	
-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--

BP 191 91006 EVRY celex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
On Jan 15, 2003 this sequence version replaced gi:20177340.  
----- Genome Center  
Center: Genoscope / Centre National de Sequencage  
Center code: GS  
Web site: http://www.genoscope.cns.fr/  
Contact: SeqRef@genoscope.cns.fr

The following sequence is oriented from the T7 to the SP6 end.  
Upstream BAC (overlapping the T7 end) : C-2240H23 (AC=AL35601.7)  
Downstream BAC (overlapping the SP6 end) : R-1101H7\_PCR1  
----- Finishing boundaries

FINISHED SEGMENT STARTS AT BASE 1  
FINISHED SEGMENT ENDS AT BASE 179128

----- Summary Statistics  
Assembly program: Phrap; version 2.0

-----  
Overall quality chart :

Range : bases  
0 :  
1 - 9 :  
10 - 19 : 2  
20 - 29 : 18  
30 - 39 : 84  
40 - 49 : 1914  
50 - 59 : 9448  
60 - 69 : 10854  
70 - 79 : 24792  
80 - 89 : 59023  
90 - 99 : 71657

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES  
source

1. 177792  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="14"  
/clone\_lib="RPCI-11"  
4329. 4475  
/note="matching EMBL:G37684  
RHdb:RH100285  
RHdb:RH81439  
dbSTS:STS57853  
Identified using the e-PCR software (G. Schuler)"  
4331. 4503  
/note="matching EMBL:AA262488  
RHdb:RH68069  
dbSTS:STS47996  
Identified using the e-PCR software (G. Schuler)"  
4479. 4599  
/note="matching EMBL:AA251552  
RHdb:RH103309  
dbSTS:STS70843  
Identified using the e-PCR software (G. Schuler)"  
9848. 10006  
/note="matching EMBL:AA069444  
RHdb:RH47658  
dbSTS:STS40717  
Identified using the e-PCR software (G. Schuler)"  
10469. 10653  
/note="matching EMBL:AA931786  
RHdb:RH98980  
dbSTS:STS68728  
Identified using the e-PCR software (G. Schuler)"  
13456. 13691  
/note="matching EMBL:Z39490  
RHdb:RH10940  
dbSTS:STS29-09

repeat\_region

Identified using the e-PCR software (G. Schuler)"  
33258^33259  
/note="TGGAATTGCCCTATATTTCAGACATCTGTATACATTAAACCATTAACAA  
G  
CGCTGCCGACAGATATCCCGTGGGAGGAGTAAACCCAGCGCACTATCGGATGCCAT  
T  
GTATATAAGCTCGAAAGGCTCTGCAAGGTTCTTTGCTGCGGTTAACCGCTCTGGTTG  
G  
CATGATACCTGATGTAGTCACGCTTTATCGTTTTCACGAAGTCTCTGCTATTTCGGTTA  
C  
CTCCGACTCCGACCGCGGCTTCTTCGGTTCAAGTCCCAACATCCCGGCGCACTGG  
C  
TCTTTCAITAGCCCGGTAGCATGAACCAATTATCCGTACGCCACTCCACTGGAGACAC  
G  
AAAGATCTTTGCCGAAGCGGTTCCACCGCTCCAGCATGAGCGTCTGTACTGTTTCA  
C  
GTTGAAGCGCGGTAGTACGCGCCAGTGCAGTGCCTCAGATCACAGCAGTCCAGC  
G  
GAACTGACACGAGTCTCTTCCTGTTATCACAGCAGAACTCGAACCCGTCAGAGCAC  
C  
TCGCTGATTGCTTTTCAAGGCCACTCTGCTGTATGTGCCCGTTTCGATGGCGGT  
A  
AGCAGGTTTTCGCTCAAGCAACAGCGCATTTCTGGCGCATGATCCGTTAAACACGTTT  
G  
ATTGATCGCAGCATACCATCAAGTTCTGCTGCTCTGCGAGCAGCGCCCATACCCGA  
C  
ATAACCATACGTTGGCAGCTCTCCGAAATACATGGTGTATACGGAAGACACATCCGTA  
T  
ATCAGTGTGACGACTGCGCGGCCATCCATCCAGTCATCGGTTCTGTGAGAAATGACG  
T  
CAACTGCGCACCGGACACCGGAGACAAACGCTGACTAAGCTTACTCCCATCCCGG  
G  
AATAAGGCGCGTGGCTATCCACTTTTTCGCGGCTCCATATTCACAGGCTTCTTTGA  
G  
AGTTCAATTTTCCATCGTTTCTTTCGCGAGCAGCGCTGGAGTTCTTTAATCTGCTTCA  
T  
CGGCGACGAAATTCAGCAGCAGCAACACTGTTCTCCGCGCGGAGCAGCAGTAAGAC  
T  
CCTTCTGCTATTTGCTTACGCGAGAAATACTGGCTGGCTGCTACACCATGTTGCC  
G  
GCAACGAGGAGACCGTCAATCCCGGTTTCAAGCTCTGCTGAAACAAATTCGATCTTTT  
C  
TGTGTGTAAGCGCTGCTGCTTCTCCGCGCCCTAAGACATCAATCATCTGTTCTCCAA  
T  
ACTAGTCTAAACACTAGTATTAAGACTATACACTTATTAAGTGAATTTGTTCTGCTGG  
A  
ATTACGCGCGCGCTA"  
/evidence=not experimental  
/insertion\_seq="IS2"

ORIGIN

Query Match 71.5%; Score 18.6; DB 9; Length 177792;  
Best Local Similarity 84.0%; Pred. No. 3e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCACTTATCCCGTTTGTACAGA 25  
|||  
Db 31985 TCACTTCTGCTGCTTGGCCAGA 31961  
|||

Search completed: April 20, 2004, 09:33:34  
Job time : 794.506 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 02:07:57 ; Search time 166.859 Seconds

(without alignments)  
661.956 Million cell updates/sec

Title: US-09-877-819b-37

Perfect score: 26

Sequence: 1 tcaacttatgcgcgtttgtacagac 26

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002s.\*

7: Geneseq2003as.\*

8: Geneseq2003bs.\*

9: Geneseq2003cs.\*

10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	26	6	AAL48201 Human HLA
2	26	100.0	255	6	AAL48219 Human HLA
3	26	100.0	267	4	AAK54418 Human hae
4	26	100.0	267	4	AAK54603 Human hae
5	26	100.0	272	4	AAK54317 Human hae
6	26	100.0	294	4	AAK54555 Human hae
7	26	100.0	294	4	AAK54805 Human hae
8	26	100.0	362	3	AAA43013 Human sec
9	26	100.0	396	2	AAV86130 EST clone
10	26	100.0	410	8	ACH49248 Human leu
11	26	100.0	436	3	AAH43818 Human sec
12	26	100.0	466	2	AAV86216 Human sec
13	26	100.0	476	3	AAC00117 Human sec
14	26	100.0	490	8	ACH49928 Human hae
15	26	100.0	576	3	AAC00116 Human sec
16	26	100.0	1202	2	AAQ25060 Human dia
17	26	100.0	1259	4	AAH31123 Human dia
18	26	100.0	1348	6	AAH18332 Human hae
19	26	100.0	14646	6	ABK64796 Human hae
20	19.6	75.4	175737	4	ABK63571 Human cDN
21	18	69.2	18	6	ABL30835 Human HLA
22	18	69.2	1380	4	ABL24141 Drosophil
23	18	69.2	3651	4	ABL24140 Drosophil

4794	69.2	18	24	ABL12632	Drosophil
6	67.7	17.6	25	ABK39244	Identific
30	67.7	17.6	26	AAK75954	Human imm
4	67.7	17.6	27	AAK75953	Human imm
304	67.7	17.6	28	AAK75952	Human imm
304	67.7	17.6	29	ABK80292	Bacillus
534	66.9	17.4	30	ADa58967	Maize hex
29	66.2	17.2	31	ADa58966	Maize hex
197	66.2	17.2	32	ADa58965	Maize hex
296	66.2	17.2	33	ADa58964	Maize hex
307	66.2	17.2	34	ADa58963	Maize hex
420	66.2	17.2	35	ADa58962	Maize hex
8	66.2	17.2	36	ADa58961	Maize hex
1524	66.2	17.2	37	ADa58960	Maize hex
7	66.2	17.2	38	ADa58959	Maize hex
1164	66.2	17.2	39	ADa58958	Maize hex
9	66.2	17.2	40	ADa58957	Maize hex
1164	66.2	17.2	41	ADa58956	Maize hex
9	66.2	17.2	42	ADa58955	Maize hex
1848	66.2	17.2	43	ADa58954	Maize hex
2006	66.2	17.2	44	ADa58953	Maize hex
9	66.2	17.2	45	ADa58952	Maize hex
2370	66.2	17.2	46	ADa58951	Maize hex
7	66.2	17.2	47	ADa58950	Maize hex
3171	66.2	17.2	48	ADa58949	Maize hex
7	66.2	17.2	49	ADa58948	Maize hex
3171	66.2	17.2	50	ADa58947	Maize hex
7	66.2	17.2	51	ADa58946	Maize hex
3171	66.2	17.2	52	ADa58945	Maize hex
7	66.2	17.2	53	ADa58944	Maize hex
486	66.2	17.2	54	ADa58943	Maize hex
1032	66.2	17.2	55	ADa58942	Maize hex
3	66.2	17.2	56	ADa58941	Maize hex
1173	66.2	17.2	57	ADa58940	Maize hex
7	66.2	17.2	58	ADa58939	Maize hex
1539	66.2	17.2	59	ADa58938	Maize hex
7	66.2	17.2	60	ADa58937	Maize hex
1863	66.2	17.2	61	ADa58936	Maize hex
3	66.2	17.2	62	ADa58935	Maize hex
2016	66.2	17.2	63	ADa58934	Maize hex
3	66.2	17.2	64	ADa58933	Maize hex
3577	66.2	17.2	65	ADa58932	Maize hex
1	66.2	17.2	66	ADa58931	Maize hex
4578	66.2	17.2	67	ADa58930	Maize hex
1	66.2	17.2	68	ADa58929	Maize hex
4578	66.2	17.2	69	ADa58928	Maize hex
2	66.2	17.2	70	ADa58927	Maize hex
4580	66.2	17.2	71	ADa58926	Maize hex
2	66.2	17.2	72	ADa58925	Maize hex
2	66.2	17.2	73	ADa58924	Maize hex
7664	66.2	17.2	74	ADa58923	Maize hex
6	66.2	17.2	75	ADa58922	Maize hex
107	66.2	17.2	76	ADa58921	Maize hex
6	66.2	17.2	77	ADa58920	Maize hex
8142	66.2	17.2	78	ADa58919	Maize hex
338	66.2	17.2	79	ADa58918	Maize hex
9	66.2	17.2	80	ADa58917	Maize hex
3	66.2	17.2	81	ADa58916	Maize hex
464	66.2	17.2	82	ADa58915	Maize hex
3	66.2	17.2	83	ADa58914	Maize hex
546	66.2	17.2	84	ADa58913	Maize hex
8	66.2	17.2	85	ADa58912	Maize hex
570	66.2	17.2	86	ADa58911	Maize hex
6	66.2	17.2	87	ADa58910	Maize hex
619	66.2	17.2	88	ADa58909	Maize hex
8	66.2	17.2	89	ADa58908	Maize hex
624	66.2	17.2	90	ADa58907	Maize hex
6	66.2	17.2	91	ADa58906	Maize hex
624	66.2	17.2	92	ADa58905	Maize hex
6	66.2	17.2	93	ADa58904	Maize hex
624	66.2	17.2	94	ADa58903	Maize hex
6	66.2	17.2	95	ADa58902	Maize hex
624	66.2	17.2	96	ADa58901	Maize hex
624	66.2	17.2	97	ADa58900	Maize hex
624	66.2	17.2	98	ADa58899	Maize hex
624	66.2	17.2	99	ADa58898	Maize hex
624	66.2	17.2	100	ADa58897	Maize hex

4794	69.2	18	24	ABL12632	Drosophil
6	67.7	17.6	25	ABK39244	Identific
30	67.7	17.6	26	AAK75954	Human imm
4	67.7	17.6	27	AAK75953	Human imm
304	67.7	17.6	28	AAK75952	Human imm
304	67.7	17.6	29	ABK80292	Bacillus
534	66.9	17.4	30	ADa58967	Maize hex
29	66.2	17.2	31	ADa58966	Maize hex
197	66.2	17.2	32	ADa58965	Maize hex
296	66.2	17.2	33	ADa58964	Maize hex
307	66.2	17.2	34	ADa58963	Maize hex
420	66.2	17.2	35	ADa58962	Maize hex
8	66.2	17.2	36	ADa58961	Maize hex
1524	66.2	17.2	37	ADa58960	Maize hex
7	66.2	17.2	38	ADa58959	Maize hex
1164	66.2	17.2	39	ADa58958	Maize hex
9	66.2	17.2	40	ADa58957	Maize hex
1164	66.2	17.2	41	ADa58956	Maize hex
9	66.2	17.2	42	ADa58955	Maize hex
1848	66.2	17.2	43	ADa58954	Maize hex
2006	66.2	17.2	44	ADa58953	Maize hex
9	66.2	17.2	45	ADa58952	Maize hex
2370	66.2	17.2	46	ADa58951	Maize hex
7	66.2	17.2	47	ADa58950	Maize hex
3171	66.2	17.2	48	ADa58949	Maize hex
7	66.2	17.2	49	ADa58948	Maize hex
3171	66.2	17.2	50	ADa58947	Maize hex
7	66.2	17.2	51	ADa58946	Maize hex
3171	66.2	17.2	52	ADa58945	Maize hex
7	66.2	17.2	53	ADa58944	Maize hex
486	66.2	17.2	54	ADa58943	Maize hex
1032	66.2	17.2	55	ADa58942	Maize hex
3	66.2	17.2	56	ADa58941	Maize hex
1173	66.2	17.2	57	ADa58940	Maize hex
7	66.2	17.2	58	ADa58939	Maize hex
1539	66.2	17.2	59	ADa58938	Maize hex
7	66.2	17.2	60	ADa58937	Maize hex
1863	66.2	17.2	61	ADa58936	Maize hex
3	66.2	17.2	62	ADa58935	Maize hex
2016	66.2	17.2	63	ADa58934	Maize hex
3	66.2	17.2	64	ADa58933	Maize hex
3577	66.2	17.2	65	ADa58932	Maize hex
1	66.2	17.2	66	ADa58931	Maize hex
4578	66.2	17.2	67	ADa58930	Maize hex
1	66.2	17.2	68	ADa58929	Maize hex
4578	66.2	17.2	69	ADa58928	Maize hex
2	66.2	17.2	70	ADa58927	Maize hex
4580	66.2	17.2	71	ADa58926	Maize hex
2	66.2	17.2	72	ADa58925	Maize hex
2	66.2	17.2	73	ADa58924	Maize hex
7664	66.2	17.2	74	ADa58923	Maize hex
6	66.2	17.2	75	ADa58922	Maize hex
107	66.2	17.2	76	ADa58921	Maize hex
6	66.2	17.2	77	ADa58920	Maize hex
8142	66.2	17.2	78	ADa58919	Maize hex
338	66.2	17.2	79	ADa58918	Maize hex
9	66.2	17.2	80	ADa58917	Maize hex
3	66.2	17.2	81	ADa58916	Maize hex
546	66.2	17.2	82	ADa58915	Maize hex
8	66.2	17.2	83	ADa58914	Maize hex
570	66.2	17.2	84	ADa58913	Maize hex
6	66.2	17.2	85	ADa58912	Maize hex
619	66.2	17.2	86	ADa58911	Maize hex
8	66.2	17.2	87	ADa58910	Maize hex
624	66.2	17.2	88	ADa58909	Maize hex
6	66.2	17.2	89	ADa58908	Maize hex
624	66.2	17.2	90	ADa58907	Maize hex
624	66.2	17.2	91	ADa58906	Maize hex
624	66.2	17.2	92	ADa58905	Maize hex
624	66.2	17.2	93	ADa58904	Maize hex
624	66.2	17.2	94	ADa58903	Maize hex
624	66.2	17.2	95	ADa58902	Maize hex
624	66.2	17.2	96	ADa58901	Maize hex

97	16.4	63.1	3521	7	ABX76282	Abx76282 lung can
98	16.4	63.1	7880	4	AAS28527	Genomic s
99	16.4	63.1	7880	4	AAS28526	Genomic s
100	16.4	63.1	21636	3	AAS5966	Human G71
101	16.4	63.1	34279	4	AAP28539	Genomic f
102	16.4	63.1	59290	3	AAP22281	BAC conta
103	16.4	63.1	96594	9	ADC85476	Human Maf
104	16.4	63.1	96595	8	ADAC2996	Human Maf
105	16.4	63.1	96595	9	ADB72734	Human Maf
106	16.4	63.1	110000	6	ABN71527_03	Continuation (4 of
107	16.4	63.1	112288	6	ABK83569	Human cDN
108	16.4	63.1	159400	6	ABQ88126	Human ost
109	16.4	63.1	202001	6	ABS52506	Human tra
110	16.2	62.3	25	8	AC162461	Human mic
111	16.2	62.3	507	9	ADBS3928	Human pro
112	16.2	62.3	1086	5	AAS92778	DNA encod
113	16.2	62.3	1548	5	AAS85698	DNA encod
114	16.2	62.3	1630	4	AAH02236	Succinivi
115	16.2	62.3	1842	4	AAS52491	E. coli D
116	16.2	62.3	1842	7	ACA32564	Prokaryot
117	16.2	62.3	2022	7	ACA54096	Prokaryot
118	16.2	62.3	2396	7	ACA20632	Prokaryot
119	16.2	62.3	2571	8	ADA31842	DNA encod
120	16.2	62.3	3177	7	ACA48237	Prokaryot
121	16.2	62.3	110000	2	AAV21209_11	Continuation (12 o
122	16.2	61.5	313	4	ABK42293	Genomic s
123	16.2	61.5	313	8	ADB60449	Connectiv
124	16.2	61.5	432	4	AAH98861	Human EST
125	16.2	61.5	561	8	ACL19944	DNA clone
126	16.2	61.5	732	4	AAH84592	E. coli g
127	16.2	61.5	732	7	ACA18612	Prokaryot
128	16.2	61.5	780	5	AAS92730	DNA encod
129	16.2	61.5	1000	8	ADB23243	Enviroeme
130	16.2	61.5	1158	2	AAZ20172	Bovine pr
131	16.2	61.5	1197	7	ACF71044	Phototrab
132	16.2	61.5	1255	4	AAH29799	S cerevis
133	16.2	61.5	2000	6	ABZ16900	Arabidops
134	16.2	61.5	2130	2	AAV54121	Human mem
135	16.2	61.5	2484	4	ABL05154	Drosophil
136	16.2	61.5	2573	3	AAH58624	Human PRO
137	16.2	61.5	2573	6	ABL88095	Human PRO
138	16.2	61.5	2573	6	ABL95584	Human ang
139	16.2	61.5	2573	9	ADD10336	Human sec
140	16.2	61.5	2573	9	ADD11296	Human sec
141	16.2	61.5	2573	9	ADD37089	Human sec
142	16.2	61.5	2573	10	ADB41297	Human sec
143	16.2	61.5	2606	4	ABL06285	Drosophil
144	16.2	61.5	2609	4	ABL21125	Drosophil
145	16.2	61.5	3203	9	ADD93582	Arabidops
146	16.2	61.5	4350	2	AAZ12252	Neisseria
147	16.2	61.5	5032	4	ABL21124	Drosophil
148	16.2	61.5	5401	4	ABL06284	Drosophil
149	16.2	61.5	6507	5	AAH66920	C glutami
150	16.2	61.5	6507	7	ACA01438	C. glucan
151	16.2	61.5	7869	4	ABL05134	Drosophil
152	16.2	61.5	9827	8	ACD19053	E. coli u
153	16.2	61.5	14974	4	ABL10974	Drosophil
154	16.2	61.5	43680	6	ABK62024	Human gen
155	16.2	61.5	5072	4	ABL10156	Drosophil
156	16.2	61.5	58175	9	ADC00292	Enterococ
157	16.2	61.5	110000	7	ACF67367_41	Continuation (42 o
158	16.2	61.5	110000	7	ACF65388_06	Continuation (7 of
159	16.2	61.5	305107	4	AAH62659	Shrimp wh
160	16.2	61.5	349880	5	AAH68530	C glutami
161	15.8	60.8	2055	8	ADA30640	DNA encod
162	15.8	60.8	2343	9	ADB3601	Human cDN
163	15.8	60.8	2691	4	AAH17791	Human cDN
164	15.8	60.8	2691	9	ADC37230	Nuclear f
165	15.8	60.8	2991	2	AAH13168	Enterococ
166	15.8	60.8	2991	6	ABS98963	Enterococ
167	15.8	60.8	3626	9	ADC37236	Nuclear f
168	15.8	60.8	3833	9	ADC37238	Nuclear f
169	15.6	60.0	277	8	ADA58955	Maize hex

ALIGNMENTS

RESULT :  
AAL48201  
ID AAL48201 standard; DNA; 26 BP.  
XX  
AC AAL48201;  
XX  
DT 01-OCT-2002 (first entry)  
XX  
DE Human HLA DPAL locus polymorphism primer sequence #2.  
XX  
KW Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;  
KW flow cytometry; human; DPAL; PCR; primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200194639-A1.  
XX  
PD 13-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US018590.  
XX  
PR 08-JUN-2000; 2000US-0210759P.  
XX  
PA (SEGC ) UNIV CALIFORNIA.  
XX  
PI White PS, Torney DC;  
XX  
DR WPI; 2002-566450/60.  
XX  
PS Identifying sequences useful as address/capture tags for flow cytometry  
PT based minisequencing, by generating tag sequences and rejecting sequences  
ET based on certain parameters e.g. sequences which form stable hairpins.  
XX  
PS Disclosure; Page 14; 35pp; English.

The present invention relates to a method of identifying sequences useful  
as address/capture tags, involving rejecting sequences having common sub-  
sequences with a sub-sequence length greater than specified number of  
bases, and sequences which can form stable hairpins and stable dimers  
from a sample of oligonucleotides, and selecting those sequences in the  
sample that would hybridize to their respective complements with a high  
degree of specificity. The method is useful for identifying a set of  
sequences useful as address/capture tags which can be used for  
multiplexed single nucleotide polymorphism (SNP) scoring in a flow  
cytometry assay. The present sequence is a primer tag described in the  
exemplification of the invention

Sequence 26 BP; 6 A; 7 C; 5 G; 8 T; 0 U; 0 Other;  
Query Match 100.0%; Score 26; DB 6; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26  
Db 1 TCAACTTATGCGCGCTTTGTACAGAC 26

```
RESULT 2
AAL48219
ID AAL48219 standard; DNA; 255 BP.
XX AC AAL48219;
XX 01-OCT-2002 (first entry)
XX DE Humar HLA DPAL exon 2 sequence.
XX
XX Flow cytometry based minisequencing; single nucleotide polymorphism; SNP;
XX flow cytometry; human; DPAL; gene; ds.
XX Homo sapiens.
XX WO200194639-A1.
XX 13-DEC-2001.
XX
XX 07-JUN-2001; 2001WO-US018590.
XX
XX 08-JUN-2000; 2000US-0210759P.
XX (REGC ) UNIV CALIFORNIA.
XX White PS, Torney DC;
XX WPI; 2002-566450/60.
XX
XX Identifying sequences useful as address/capture tags for flow cytometry
XX based minisequencing, by generating tag sequences and rejecting sequences
XX based on certain parameters e.g. sequences which form stable hairpins.
XX
XX Disclosure; Fig 4; 35pp; English.
XX
XX The present invention relates to a method of identifying sequences useful
XX as address/capture tags, involving rejecting sequences having common sub-
XX sequences with a sub-sequence length greater than specified number of
XX bases, and sequences which can form stable hairpins and stable dimers
XX from a sample of oligonucleotides, and selecting those sequences in the
XX sample that would hybridise to their respective complements with a high
XX degree of specificity. The method is useful for identifying a set of
XX sequences useful as address/capture tags which can be used for
XX multiplexed single nucleotide polymorphism (SNP) scoring in a flow
XX cytometry assay. The present sequence is a fragment of the human DPAL
XX gene described in the exemplification of the invention
XX
XX Sequence 255 BP; 67 A; 57 C; 66 G; 65 T; 0 U; 0 Other;
Query Match 100.0%; Score 26; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTTGTCAGAC 26
Db 19 TCAACTTATGCGCGTTTGTCAGAC 44

RESULT 3
AAK54418
ID AAK54418 standard; cDNA; 267 BP.
XX AC AAK54418;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #143.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
```

```
OS Homo sapiens.
XX WO200164886-A2.
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0222903P.
XX 04-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 371; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of haematological malignancies. The
XX present sequence is the coding sequence of a human haematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of haematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Haematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
Query Match 100.0%; Score 26; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTTGTCAGAC 26
Db 140 TCAACTTATGCGCGTTTGTCAGAC 165

RESULT 4
AAK54603
ID AAK54603 standard; cDNA; 267 BP.
XX AC AAK54603;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human haematological malignancy-related antigen coding sequence #328.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX WO200164886-A2.
XX
```

us-09-877-819b-37.rng

Wed Apr 21 10:46:47 2004

```

PF 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0222903P.
XX 04-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 418; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 26; DB 4; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 0.0036;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TCAACTTATGCGCGTTGTACAGAC 26
DB 140 TCAACTTATGCGCGTTGTACAGAC 165
XX
XX RESULT 5
XX AAK54317
XX ID AAK54317 standard; cDNA; 272 BP.
XX
XX AC AAK54317;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human hematological malignancy-related antigen coding sequence #42.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX hematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO200164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX 17-MAR-2000; 2000US-0190479P.
XX 27-APR-2000; 2000US-0200545P.
XX 28-APR-2000; 2000US-0200303P.
XX 28-APR-2000; 2000US-0200779P.
XX 01-MAY-2000; 2000US-0200999P.
XX 04-MAY-2000; 2000US-0202084P.
XX 22-MAY-2000; 2000US-0206201P.
XX 14-JUL-2000; 2000US-0218950P.
XX 03-AUG-2000; 2000US-0222903P.
XX 04-AUG-2000; 2000US-0223416P.
XX 07-AUG-2000; 2000US-0223378P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 418; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 267 BP; 67 A; 59 C; 73 G; 68 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 26; DB 4; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 0.0036;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TCAACTTATGCGCGTTGTACAGAC 26
DB 140 TCAACTTATGCGCGTTGTACAGAC 165
XX
XX RESULT 6
XX AAK54555/c
XX ID AAK54555 standard; cDNA; 294 BP.
XX
XX AC AAK54555;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human hematological malignancy-related antigen coding sequence #280.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX hematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO200164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186126P.
XX

```

us-09-877-819b-37.rng

Wed Apr 21 10:46:47 2004

```

PR 17-MAR-2000; 2000US-0190479P.
PR 27-APR-2000; 2000US-0200545P.
PR 28-APR-2000; 2000US-0200303P.
PR 28-APR-2000; 2000US-0200779P.
PR 01-MAY-2000; 2000US-0200999P.
PR 04-MAY-2000; 2000US-020084P.
PR 22-MAY-2000; 2000US-0218950P.
PR 14-JUL-2000; 2000US-0222903P.
PR 03-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J;
XX PI
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 406; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 26; DB 4; Length 294;
XX Best Local Similarity 100.0%; Pred. No. 0.0036;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCACCTTATGCGCGTTTGTACAGAC 26
DB 128 TCACCTTATGCGCGTTTGTACAGAC 103
XX
RESULT 7
AAK54805/c
ID AAK54805 standard; cDNA; 294 BP.
XX
XX AAK54805;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human hematological malignancy-related antigen coding sequence #530.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX hematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Homo sapiens.
XX
XX WO200164886-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001; 2001WO-US007272.
XX
XX 01-MAR-2000; 2000US-0186125P.
XX
XX 17-MAR-2000; 2000US-0190479P.
XX
XX 27-APR-2000; 2000US-0200545P.
XX
XX 28-APR-2000; 2000US-0200303P.

```

```

PR 28-APR-2000; 2000US-0200779P.
PR 01-MAY-2000; 2000US-0200999P.
PR 04-MAY-2000; 2000US-020084P.
PR 22-MAY-2000; 2000US-0206201P.
PR 14-JUL-2000; 2000US-0218950P.
PR 03-AUG-2000; 2000US-0222903P.
PR 04-AUG-2000; 2000US-0223416P.
PR 07-AUG-2000; 2000US-0223378P.
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J;
XX PI
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
XX e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 469; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/3
XX cell non-Hodgkin's lymphoma
XX
XX Sequence 294 BP; 79 A; 77 C; 68 G; 70 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 26; DB 4; Length 294;
XX Best Local Similarity 100.0%; Pred. No. 0.0036;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCACCTTATGCGCGTTTGTACAGAC 26
DB 128 TCACCTTATGCGCGTTTGTACAGAC 103
XX
RESULT 8
AAA43013
ID AAA43013 standard; cDNA; 362 BP.
XX
XX AAA43013;
XX
XX 21-AUG-2000 (first entry)
XX
XX Human secreted expressed sequence tag SEQ ID NO:1753.
XX
XX Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;
XX expressed sequence tag; ESI; probe; chemotactic; proliferative;
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
XX anticancer; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
XX autoimmune disorder; multiple sclerosis; allergic condition;
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
XX central nervous system disorder; Alzheimer's disease; stroke;
XX Parkinson's disease; Huntington's disease; coagulation disorder;
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
XX infection; depression; psoriasis; ss.
XX
XX Homo sapiens.
XX
XX WO2000021990-A1.
XX
XX

```

PD 20-APR-2000.  
 XX 15-OCT-1999; 99WO-US024205.  
 PF 15-OCT-1998; 98US-0104435P.  
 PR 15-OCT-1998; 98US-0104435P.  
 XX (GEMY) GENETICS INST INC.  
 PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 XX Merberg D, Treacy M;  
 PI WPI; 2000-317937/27.  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 XX expressed sequence tags (sESTs), useful for treating various disorders  
 XX such as autoimmune, infectious, and central nervous system disorders.  
 XX Claim 1; Page 526; 619pp; English.  
 PS  
 XX AAA41261 to AAA43419 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, xenopus and rat tissue  
 CC sources. The sESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; anti-inflammatory; cytostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antiparkinsonian; antipsoriatic;  
 CC osteoprotective; neuroprotective; nootropic; antidepressant. The sESTs can be  
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention  
 XX Sequence 362 BP; 86 A; 86 C; 99 G; 91 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 26; DB 3; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0038;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCAACTTATGCGCGTTGTACAGAC 26  
 DB 144 TCAACTTATGCGCGTTGTACAGAC 169  
 RESULT 9  
 AAV86130  
 ID AAV86130 standard; cDNA; 396 BP.  
 XX  
 AC AAV86130;  
 XX  
 DT 27-APR-1999 (first entry)  
 XX EST clone H45.  
 DE  
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 XX Homo sapiens.  
 OS  
 XX WO9845435-A2.  
 PN  
 XX 15-OCT-1998.  
 PD

XX 10-APR-1998; 98WO-US006954.  
 XX 10-APR-1997; 97US-00835913.  
 XX (GEMY) GENETICS INST INC.  
 XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 XX Spaulding V, Agostino MJ;  
 XX WPI; 1999-070076/06.  
 XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 XX pituitary, retina and colon cDNA libraries.  
 XX Claim 1; Page 131; 633pp; English.  
 XX This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX Sequence 396 BP; 100 A; 97 C; 101 G; 98 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 26; DB 2; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.0038;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCAACTTATGCGCGTTGTACAGAC 26  
 DB 163 TCAACTTATGCGCGTTGTACAGAC 198  
 RESULT 10  
 ACH49248  
 ID ACH49248 standard; cDNA; 410 BP.  
 XX  
 AC ACH49248;  
 XX  
 DT 13-OCT-2003 (first entry)  
 XX Human leukocyte cDNA #842.  
 DE  
 XX Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 XX Homo sapiens.  
 OS  
 XX US2003073623-A1.  
 PN  
 XX 17-APR-2003.  
 PD  
 XX 30-JUL-2001; 2001US-00918995.  
 PF  
 XX 30-JUL-2001; 2001US-00918995.  
 PR  
 XX (DRMA/) DRMANAC R T.  
 XX (LABA/) LABAT I.  
 PA (STAC/) STACHE-CRAIN B.  
 PA (DICK/) DICKSON M C.  
 PA (JONE/) JONES L W.  
 XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
 PI

XX DR WPI; 2003-615964/58.

XX PT New polynucleotide sequences obtained from various cDNA libraries, useful

XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene

XX PT mapping, in the recombinant production of protein, or in generating

XX PT antisense DNA or RNA.

XX PS Claim 1; SEQ ID NO 36460; 44pp; English.

XX PT The invention relates to an isolated polynucleotide comprising any one of

XX CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was

XX CC determined by the technique of SBH (sequencing by hybridisation). Also

XX CC included is a purified polypeptide comprising a sequence corresponding to

XX CC a reading frame of the novel polynucleotide. The nucleic acid sequences

XX CC are useful in diagnostics as expressed sequence tags (EST) for

XX CC identifying expressed genes or for physical mapping of the human genome,

XX CC in forensics, in assessing biodiversity, or in identifying mutations

XX CC responsible for genetic disorders and other traits. The nucleotide

XX CC sequences are also useful as hybridisation probes, as oligomers for PCR,

XX CC for chromosome and gene mapping, in the recombinant production of

XX CC protein, or in generating antibodies specific for it. The present sequence

XX CC is useful for generating antibodies specific for it. The present sequence data

XX CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data

XX CC for this patent did not form part of the printed specification, but was

XX CC obtained in electronic format directly from USPTO at

XX CC seqdata.uspto.gov/sequence.html?docID=20030073623

XX SQ Sequence 410 BP; 99 A; 105 C; 101 G; 105 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 8; Length 410;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26

DB 163 TCAACTTATGCGCGTTGTACAGAC 188

RESULT 11

AAA43818

ID AAA43818 standard; cDNA; 436 BP.

XX AC AAA43818;

XX DT 21-AUG-2000 (first entry)

XX DE Mouse secreted expressed sequence tag SEQ ID NO:393.

XX KW Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;

XX KW expressed sequence tag; sEST; probe; chemotactic; proliferative;

XX KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;

XX KW thrombolytic; antiinflammatory; cyostatic; antibacterial; antifungal;

XX KW antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;

XX KW antitumor; osteopathic; neuroprotective; neurotropic; antipsoriatic;

XX KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;

XX KW autoimmune disorder; multiple sclerosis; allergic condition;

XX KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;

XX KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;

XX KW central nervous system disorder; Alzheimer's disease; stroke;

XX KW Parkinson's disease; Huntington's disease; coagulation disorder;

XX KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;

XX KW infection; depression; psoriasis; ss.

XX OS Mus musculus.

XX PN WO200021991-A1.

XX PD 20-APR-2000.

XX PF 15-OCT-1999; 99WO-US024206.

XX PR 15-OCT-1999; 98US-0104436P.

XX BA (GENY) GENETICS INST INC.

XX PT Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX PT Merberg D, Treacy M, Bowman MR;

XX DR WPI; 2000-317938/27.

XX PT Isolated polynucleotides, and encoded proteins, comprising secreted

XX PT expressed sequence tags (sESTs), useful for treating various disorders

XX PT such as autoimmune, infectious, and central nervous system disorders.

XX PS Claim 1; Page 306; 803pp; English.

XX CC AAA43426 to AAA45925 represent specifically claimed secreted expressed

XX CC sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue

XX CC sources. The sESTs can have a range of activities depending on the

XX CC tissues they were isolated from. The activities include: chemotactic;

XX CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;

XX CC haemostatic; thrombolytic; antiinflammatory; cyostatic; antibacterial;

XX CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnerary; antitumor;

XX CC cerebroprotective; anticonvulsant; antidepressant. The sESTs can be

XX CC used for gene therapy and in vaccines. The sESTs are useful as probes for

XX CC the identification and isolation of full-length cDNAs and genomic DNA

XX CC molecules which correspond to the sESTs. Proteins encoded by the sESTs

XX CC are useful in assays for determining biological activity and raising

XX CC antibodies. They may be useful for treatment of autoimmune disorders

XX CC (multiple sclerosis, insulin dependent diabetes), allergic conditions

XX CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,

XX CC osteoporosis, osteoarthritis, central nervous system disorders

XX CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation

XX CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's

XX CC disease), tumours, bacterial, fungal or viral infections, depression and

XX CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given

XX CC in the exemplification of the present invention

XX SQ Sequence 436 BP; 108 A; 113 C; 106 G; 109 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 3; Length 436;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26

DB 185 TCAACTTATGCGCGTTGTACAGAC 210

RESULT 12

AAV86216

ID AAV86216 standard; cDNA; 466 BP.

XX AC AAV86216;

XX DT 27-APR-1999 (first entry)

XX DE EST clone O57.

XX KW Expressed sequence tag; secreted protein; haematopoiesis regulator;

XX KW tissue growth; activin; inhibin; tumour invasion suppressor; sEST; human;

XX KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;

XX KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX OS Homo sapiens.

XX PN WO9845435-A2.

XX PD 15-OCT-1998.

XX PF 10-APR-1998; 98WO-US006954.

XX PR 10-APR-1997; 97US-00835913.

XX XX

(GEM) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
Spaulding V, Agostino MJ;  
WPI; 1999-070076/06.

New polynucleotides encoding human secreted proteins - derived from e.g.  
human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
pituitary, retina and colon cDNA libraries.

Claim 1; Page 160-161; 633pp; English.

This sequence represents an expressed sequence tag (EST), and is a  
polynucleotide of the invention. The polynucleotides of the invention are  
all secreted EST sequences isolated from a variety of human tissue  
sources. The EST sequences and proteins encoded by them are predicted to  
have useful biological activities which would make them suitable for  
treating, preventing or ameliorating medical conditions in humans and  
animals, although no supporting data is given. Suggested activities  
include nutritional activity, immune stimulating or suppressing activity,  
haematopoiesis regulating activity, tissue growth activity,  
activin/inhibin activity, chemotactic/chemokinetic activity,  
and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
activity. The EST sequences are also stated to be useful for gene therapy  
activity.

Sequence 466 BP; 110 A; 122 C; 122 G; 112 T; 0 U; 0 Other;  
Query Match 100.0%; Score 26; DB 2; Length 466;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCGCGTTGTACAGAC 26  
131 TCAACTTATGCGCGTTGTACAGAC 156

RESULT 13  
AAC00117  
ID AAC00117 standard; cDNA; 476 BP.

AAC00117;  
06-OCT-2000 (first entry)  
Human secreted protein 5' EST, SEQ ID NO: 115.  
Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
gene therapy; chromosome mapping; ss.  
Homo sapiens.  
EP1033401-A2.  
06-SEP-2000.  
21-FEB-2000; 2000EP-00200610.  
26-FEB-1999; 99US-0122487P.  
(GSET) GENSET.  
Dumas Milne Edwards J, Duclert A, Giordano J;  
WPI; 2000-500381/45.  
P-PSDB; ABG00111.  
New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX Claim 1; SEQ ID NO 115; 71pp + Sequence Listing; English.

The present sequence is one of a large number of 5' ESTs derived from  
mRNAs encoding secreted proteins. An ORF has been identified within the  
sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
derived from 30 different tissues. EST sequences usually correspond  
mainly to the 3' untranslated region (UTR) of the mRNA because they are not  
often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
well suited for isolating cDNA sequences derived from the 5' ends of  
mRNAs and even in those cases where longer cDNA sequences have been  
obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
mRNAs with intact 5' ends and can therefore be used to obtain full length  
cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
gene therapy and chromosome mapping procedures. They are used to obtain  
upstream regulatory sequences and to design expression and secretion  
vectors

Sequence 476 BP; 110 A; 125 C; 120 G; 121 T; 0 U; 0 Other;  
Query Match 100.0%; Score 26; DB 3; Length 476;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAACTTATGCGCGTTGTACAGAC 26  
206 TCAACTTATGCGCGTTGTACAGAC 231

RESULT 14  
ACH49928  
ID ACH49928 standard; cDNA; 490 BP.

ACH49928;  
13-OCT-2003 (first entry)  
Human leukocyte cDNA #1522.  
Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
genome mapping; biodiversity; genetic disorder.  
Homo sapiens.  
US2003073623-A1.  
17-APR-2003.  
30-JUL-2001; 2001US-00918995.  
30-JUL-2001; 2001US-00918995.  
(DRMA/) DRMANAC R T.  
(LAEA/) LABAT I.  
(STAC/) STACHE-CRAIN B.  
(DICK/) DICKSON M C.  
(JONE/) JONES L W.  
Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
WPI; 2003-615964/58.  
New polynucleotide sequences obtained from various cDNA libraries, useful  
as hybridization probes, as oligomers for PCR, for chromosome and gene  
mapping, in the recombinant production of protein, or in generating  
antisense DNA or RNA.  
Claim 1; SEQ ID NO 37140; 44pp; English.  
The invention relates to an isolated polynucleotide comprising any one of  
38043 cDNA sequences, appearing as ACH2789-ACH50831, whose sequence was  
determined by the technique of SBH (sequencing by hybridisation). Also  
included is a purified polypeptide comprising a sequence corresponding to  
a reading frame of the novel polynucleotide. The nucleic acid sequences  
are useful in diagnostics as expressed sequence tags (EST) for



us-09-877-819b-37.rng

Wed Apr 21 10:46:47 2004

identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversities, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=200300073623

Sequence 430 BP; 112 A; 130 C; 125 G; 121 T; 0 U; 2 Other;

Query Match 100.0%; Score 26; DB 8; Length 490;  
Best Local Similarity 100.0%; Pred. No. 0.004; 0; Indels 0; Gaps 0;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACTTATGCGCGTTGTACAGAC 26  
|||||  
Db 236 TCACTTATGCGCGTTGTACAGAC 261  
|||||

RESULT 15  
AAC00116  
ID AAC00116 standard; cDNA; 576 BP.  
XX AC AAC00116;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein 5' EST, SEQ ID NO: 114.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX KW gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PP 21-FEB-2000; 2000EP-00200610.  
XX PR 26-FEB-1999; 99US-0122487P.  
XX PA (GEST) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX P-PSDB; AAG00110.  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX Claim 1; SEQ ID NO 114; 71pp + Sequence Listing; English.  
XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors

XX SQ Sequence 576 BP; 131 A; 158 C; 141 G; 145 T; 0 U; 1 Other;

Query Match 100.0%; Score 26; DB 3; Length 576;  
Best Local Similarity 100.0%; Pred. No. 0.0041;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACTTATGCGCGTTGTACAGAC 26  
|||||  
Db 306 TCACTTATGCGCGTTGTACAGAC 331  
|||||

RESULT 16  
AAQ25060  
ID AAQ25060 standard; DNA; 1202 BP.  
XX AC AAQ25060;  
XX DT 15-NOV-1992 (first entry)  
XX DE pSBalpha-318 clone.  
XX KW RFLP; probe p29G8; beta-lymphoblastoid cell line; LG2; T5-1; 6.3.6;  
XX KW transplant; transfection; paternity; ss.  
XX OS Synthetic.  
XX PN US5110920-A.  
XX PD 05-MAY-1992.  
XX PF 05-DEC-1984; 84US-00678255.  
XX PR 22-JAN-1982; 82US-00341902.  
XX PR 07-JAN-1983; 83US-00456373.  
XX PR 30-AUG-1988; 88US-00238619.  
XX PA (CETU) CETUS CORP.  
XX PI Erlich HA;  
XX WPI; 1992-175244/21.  
XX New DNA probes specific to single class II HLA locus - useful in HLA  
PT typing e.g. to evaluate paternity and transplant or transfection  
PT compatibility and to diagnose disease susceptibility.  
XX Disclosure; Page 11; 21pp; English.

The sequence given is a pSBalpha-318 clone which was derived from a beta-lymphoblastoid cell line LG2 cDNA library using a probe designated p29G8. This probe bound to sequences distinct from those which lead to the elucidation of HLA-Dp34 (see also AAQ25059). A genomic blot pattern with DNA from the cell lines T5-1 and its HLA hemizygous derivative 6.3.6 indicates that the p29G8 locus maps within the HLA region. p29G8 has been found to be a HLA-SBalpha clone and could be used to isolate the given sequence. p29G8 could be useful in HLA typing based on RFLPs. It can be utilized in paternity disputes or for determining transplant or transfection compatibility. It can also be used to make disease correlations to CC diagnose diseases or predict susceptibility to diseases

Sequence 1202 BP; 289 A; 336 C; 277 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 2; Length 1202;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACTTATGCGCGTTGTACAGAC 26  
|||||  
Db 191 TCACTTATGCGCGTTGTACAGAC 216  
|||||

RESULT 17

AAS31123  
ID AAS31123 standard; cDNA; 1259 BP.

AC AAS31123;

DT 04-DEC-2001 (first entry)

XX Human diagnostic and therapeutic polynucleotide (DITHP) #138.

XX Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
KW respiratory disorder; ss.

OS Homo sapiens.

XX WO200162927-A2.

PN 30-AUG-2001.

PD 21-FEB-2001; 2001WO-US006059.

PF 24-FEB-2000; 2000US-0184693P.

XX 24-FEB-2000; 2000US-0184697P.

PR 24-FEB-2000; 2000US-0184698P.

PR 24-FEB-2000; 2000US-0184768P.

PR 24-FEB-2000; 2000US-0184769P.

PR 24-FEB-2000; 2000US-0184770P.

PR 24-FEB-2000; 2000US-0184771P.

PR 24-FEB-2000; 2000US-0184772P.

PR 24-FEB-2000; 2000US-0184773P.

PR 24-FEB-2000; 2000US-0184774P.

PR 24-FEB-2000; 2000US-0184776P.

PR 24-FEB-2000; 2000US-0184777P.

PR 24-FEB-2000; 2000US-0184779P.

PR 24-FEB-2000; 2000US-0184813P.

PR 24-FEB-2000; 2000US-0184837P.

PR 24-FEB-2000; 2000US-0184841P.

PR 24-FEB-2000; 2000US-0185213P.

PR 24-FEB-2000; 2000US-0185216P.

PR 12-MAY-2000; 2000US-0203785P.

PR 15-MAY-2000; 2000US-0204226P.

PR 16-MAY-2000; 2000US-0204525P.

PR 16-MAY-2000; 2000US-0204821P.

PR 16-MAY-2000; 2000US-0204908P.

PR 16-MAY-2000; 2000US-0205232P.

PR 17-MAY-2000; 2000US-0204815P.

PR 17-MAY-2000; 2000US-0204863P.

PR 17-MAY-2000; 2000US-0205221P.

PR 17-MAY-2000; 2000US-0205285P.

PR 17-MAY-2000; 2000US-0205286P.

PR 17-MAY-2000; 2000US-0205287P.

PR 17-MAY-2000; 2000US-0205323P.

PR 17-MAY-2000; 2000US-0205324P.

(INCY-) INCYTE GENOMICS INC.

XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
PI Chen A, D'sa SA, Amsbey S, Dahl CR, Dam TC, Daniels SE, Dufour GE;  
PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TF;  
PI Roseberry AM, Rosen BH, Russo PD, Stockdreher TK, Daffo A;  
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
XX WPI: 2001-502867/55.  
DR P-PSDB; AAU19552.

XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
XX enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
PT Claim 1; Page 361; 522pp; English.

XX The invention relates to polynucleotides (I) encoding diagnostic and

CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
CC proteins involved in growth and development and receptors. (I) and (II)  
CC may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate DITHP expression. For example, (I) and (II)  
CC may be used to treat disorders associated with decreased polypeptide  
CC expression by rectifying mutations or deletions in a patient's genome,  
CC that affect the activity of the DITHPs, by expressing inactive proteins  
CC or supplementing the patient's own production of them. (I) and (II) may  
CC be used to treat diseases, for example, cell proliferative disorder,  
CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
CC (II) may be used to produce the DITHPs, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the protein. (I) and  
CC its complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids in  
CC samples, and therefore which patients may be in need of restorative  
CC therapy. (II) may also be used as antigens in the production of  
CC antibodies against DITHPs and in assays to identify modulators of DITHP  
CC expression and activity. The anti-DITHP antibodies and antagonists may  
CC also be used to down regulate expression and activity. The anti-DITHP  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of DITHPs in samples (e.g. by enzyme-linked immunosorbent assay  
CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic  
CC (DITHP) polynucleotides of the invention

XX Sequence 1259 BP; 293 A; 350 C; 300 G; 316 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 4; Length 1259;

Best Local Similarity 100.0%; Pred. No. 0.0048; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26

Db 326 TCAACTTATGCGCGTTGTACAGAC 351

RESULT 18

AAAF18332  
ID AAF18332 standard; DNA; 1348 BP.

XX AAF18332;

XX 14-MAR-2001 (first entry)

XX Lung cancer associated polynucleotide sequence SEQ ID 351.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
XX cardioactive; immunomodulatory; muscular active; vulnery;  
XX gastrointestinal; nephrotropic; antineoplastic; gynecological;  
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
XX proliferative disorder; wound healing; infectious disease; ds.

OS Homo sapiens.

XX WO2000055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005918.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI: 2000-587514/55.

XX P-PSDB; AAS58456.

XX Lung cancer associated gene sequences, referred to as lung cancer  
XX antigens, useful for treatment, prevention, and diagnosis of disorders  
XX such as lung cancer.

	from patients exhibiting different clinical states of prostate
	hyperplasia as compared to normal prostate tissue. (I) comprises
	detecting the expression levels of one or more genes in prostatic cells
	from the subject that are differentially regulated compared to normal
	prostatic cells. (II) comprises preparing a first gene expression profile
	of BPH cells or BPH-like cell population, exposing the cells to the
	agent, preparing a second gene expression profile of the agent exposed
	cells, and comparing the first and second gene expression profiles. (I)
	is useful for diagnosing the onset or progression of BPH. (II) is useful
	for identifying an agent that modulates the onset or progression of BPH.
	The methods are useful to present information identifying the expression
	level in a tissue or cells, by comparing the expression level of genes
	expressed in the specification in the tissue or cells to the level of
	expression of gene in the database, and displaying the expression levels
	of at least one gene in the tissue or cell sample compared to the
	expression level in BPH Agents using (II) are useful for treating BPH or
	prostate cancer. ABK64106-ABK64860 represent human benign prostatic
	hyperplasia gene sequences of the invention
	Sequence 14646 BP; 4552 A; 2957 C; 2741 G; 4395 T; 0 U; 1 Other;
	Query Match            100.0%; Score 26; DB 6; Length 14646;
	Best Local Similarity    100.0%; Pred. No. 0.0079;
	Matches     26; Conservative     0; Mismatches     0; Indels     0; Gaps     0;
	1 TCAACTTATGCGCGTTGTTCACAGC 26
	6345 TCACCTATTGCCGGTTGTTCACAGC 6370
	RESULT 20
	ABX83571
	ID ASK83571 standard; cDNA; 175737 BP.
	XX AC ABK83571;
	DT
	XT 14-AUG-2002 (first entry)
	XX Human cDNA differentially expressed in granulocytic cells #142.
	DE
	XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
	KW viral infection; parasitic infection; protozoal infection;
	KW fungal infection; sterile inflammatory disease; psoriasis;
	KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
	KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
	KW adult respiratory distress syndrome; inflammatory bowel disease;
	XW Crohn's disease; ulcerative colitis; periodontal disease;
	KW granulocyte activation; chronic inflammation; allergy.
	XX
	OS Homo sapiens.
	XX WO200228999-A2.
	PN
	XX 11-APR-2002.
	PD
	XX 03-OCT-2001; 2001WO-US030821.
	PF
	XX 03-OCT-2000; 2000US-0237189P.
	PP
	(GENE-) GENE LOGIC INC.
	XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
	PA
	XX WPI; 2002-435328/46.
	PB
	DR
	XX Detecting granulocyte activation by detecting differential expression of
	PT genes associated with granulocyte activation, which serves as diagnostic
	PT markers that is useful for monitoring disease states and drug toxicity.
	XX Claim 1; SEQ ID NO 142; 114pp; English.
	PS The invention relates to detecting (M1) granulocyte (GC) activation
	XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
	CC
	CC

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Wed Apr 21 10:46:47 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:29:23 ; Search time 35.3294 Seconds

(without alignments)

408.405 Million cell updates/sec

Title: US-09-877-819B-37

Perfect score: 26

Sequence: 1 tcaactatcgccggtttgacagac 26

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*

2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*

3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*

4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*

5: /cgn2\_6/prodata/2/ina/6C COMB.seq.\*

6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	65.4	5240	4	US-09-171-337A-2
2	17	65.4	5240	4	US-09-631-022-2
3	17	65.4	5897	4	US-08-956-171E-320
4	16.9	64.6	1032	4	US-09-328-352-1688
5	16.6	63.8	4580	2	US-08-674-351-1
6	16.4	63.1	636	4	US-09-702-705-1668
7	16.4	63.1	636	4	US-09-736-457-1668
8	16.4	63.1	636	4	US-09-614-124B-1668
9	16.4	63.1	636	4	US-09-671-325-1668
10	16.4	63.1	1633	4	US-09-702-705-316
11	16.4	63.1	1633	4	US-09-736-457-316
12	16.4	63.1	1633	4	US-09-736-457-788
13	16.4	63.1	1633	4	US-09-614-124B-316
14	16.4	63.1	1633	4	US-09-614-124B-788
15	16.4	63.1	1633	4	US-09-671-325-316
16	16.4	63.1	1633	4	US-09-671-325-788
17	16.4	63.1	1633	4	US-09-589-184-316
18	16.4	63.1	1633	4	US-09-589-184-788
19	16.4	63.1	34279	4	US-09-596-002-26
20	16.4	63.1	202001	4	US-09-734-674-3
21	16.4	63.1	2571	4	US-09-328-352-3129
22	16.2	62.3	1664976	4	US-08-916-421B-1
23	16.2	61.5	451	4	US-09-621-976-3237
24	16	61.5	732	4	US-09-711-164-221-1
25	16	61.5	9827	4	US-09-453-702B-66
26	16	61.5	9827	4	US-09-134-000C-948
27	15.8	60.8	1323	4	US-09-134-000C-948

28	15.8	60.8	2055	4	US-09-328-352-1927	Sequence 1927, Ap
29	15.6	60.0	1545	4	US-09-540-236-347	Sequence 347, Ap
30	15.6	60.0	2023	4	US-08-622-191-3	Sequence 3, Appl
31	15.6	60.0	10709	4	US-09-596-002-9	Sequence 9, Appl
32	15.6	60.0	1481	4	US-09-453-702B-254	Sequence 254, Ap
33	15.6	60.0	14854	4	US-08-961-527-106	Sequence 106, Ap
34	15.4	59.2	339	1	US-08-318-970B-31	Sequence 31, Appl
35	15.4	59.2	339	1	US-08-318-970B-40	Sequence 40, Appl
36	15.4	59.2	402	1	US-08-318-970B-41	Sequence 41, Appl
37	15.4	59.2	435	4	US-09-096-731A-3	Sequence 3, Appl
38	15.4	59.2	446	4	US-09-621-976-13556	Sequence 13556, A
39	15.4	59.2	1125	4	US-09-543-681A-3439	Sequence 3439, Ap
40	15.4	59.2	1391	3	US-08-956-139-3	Sequence 3, Appl
41	15.4	59.2	1391	3	US-08-956-139-3	Sequence 3, Appl
42	15.4	59.2	1391	3	US-09-215-649A-3	Sequence 3, Appl
43	15.4	59.2	1391	4	US-09-577-800-3	Sequence 3, Appl
44	15.4	59.2	1391	4	US-09-577-800-3	Sequence 3, Appl
45	15.4	59.2	1391	4	US-09-466-496-3	Sequence 3, Appl
46	15.4	59.2	1391	4	US-08-871-856-3	Sequence 3, Appl
47	15.4	59.2	1391	4	US-08-871-291-3	Sequence 3, Appl
48	15.4	59.2	1391	4	US-09-877-650-3	Sequence 3, Appl
49	15.4	59.2	1590	4	US-09-023-655-762	Sequence 762, Ap
50	15.4	59.2	1728	4	US-09-096-731A-1	Sequence 191, Ap
51	15.4	59.2	3001	4	US-09-539-333D-191	Sequence 4156, Ap
52	15.4	59.2	3024	4	US-09-489-039A-4156	Sequence 1, Appl
53	15.4	59.2	3115	3	US-08-996-139-1	Sequence 1, Appl
54	15.4	59.2	3115	3	US-08-996-659-1	Sequence 1, Appl
55	15.4	59.2	3115	3	US-09-215-649A-1	Sequence 1, Appl
56	15.4	59.2	3115	4	US-09-577-800-1	Sequence 1, Appl
57	15.4	59.2	3115	4	US-09-466-496-1	Sequence 1, Appl
58	15.4	59.2	3115	4	US-09-871-856-1	Sequence 1, Appl
59	15.4	59.2	3115	4	US-09-871-291-1	Sequence 1, Appl
60	15.4	59.2	3115	4	US-09-877-650-1	Sequence 1, Appl
61	15.4	59.2	3136	3	US-08-996-139-5	Sequence 5, Appl
62	15.4	59.2	3136	3	US-09-435-296-3	Sequence 3, Appl
63	15.4	59.2	3136	3	US-08-995-659-5	Sequence 5, Appl
64	15.4	59.2	3136	3	US-09-215-649A-5	Sequence 5, Appl
65	15.4	59.2	3136	3	US-09-577-800-5	Sequence 5, Appl
66	15.4	59.2	3136	4	US-09-577-800-5	Sequence 5, Appl
67	15.4	59.2	3136	4	US-09-466-496-5	Sequence 5, Appl
68	15.4	59.2	3136	4	US-09-871-856-5	Sequence 5, Appl
69	15.4	59.2	3136	4	US-09-871-291-5	Sequence 5, Appl
70	15.4	59.2	3136	4	US-09-877-650-5	Sequence 5, Appl
71	15.4	59.2	3377	6	5198345-16	Patent No. 5198345
72	15.4	59.2	5531	4	US-09-620-312D-619	Sequence 619, Ap
73	15.4	59.2	5531	4	US-08-965-762-1	Sequence 1, Appl
74	15.4	59.2	5596	3	US-09-911-927-1	Sequence 1, Appl
75	15.4	59.2	5596	4	US-09-911-927-3	Sequence 3, Appl
76	15.4	59.2	5596	4	US-09-911-882-1	Sequence 1, Appl
77	15.4	59.2	5596	4	US-09-911-882-3	Sequence 3, Appl
78	15.4	59.2	5596	4	US-09-911-888-3	Sequence 3, Appl
79	15.4	59.2	5596	4	US-09-911-888-3	Sequence 3, Appl
80	15.4	59.2	1664976	4	US-08-916-421B-1	Sequence 1, Appl
81	15.4	59.2	4160	4	US-09-134-218-1	Sequence 1, Appl
82	15.2	58.5	4724	1	US-08-404-665-3	Sequence 3, Appl
83	15.2	58.5	4724	1	US-08-404-671-3	Sequence 3, Appl
84	15.2	58.5	4724	1	US-08-404-781-3	Sequence 3, Appl
85	15.2	58.5	1230025	4	US-09-198-452A-1	Sequence 1, Appl
86	15.2	58.5	155	4	US-09-621-976-9826	Sequence 9826, Ap
87	15	57.7	275	4	US-09-313-294A-3799	Sequence 3799, Ap
88	15	57.7	300	4	US-09-489-039A-3237	Sequence 3237, Ap
89	15	57.7	471	4	US-09-621-975-15417	Sequence 15417, A
90	15	57.7	582	1	US-08-446-918A-5	Sequence 5, Appl
91	15	57.7	582	1	US-08-580-806-5	Sequence 5, Appl
92	15	57.7	657	5	PCR-US95-14639-1	Sequence 1, Appl
93	15	57.7	705	1	US-08-152-456A-1	Sequence 1, Appl
94	15	57.7	705	1	US-08-440-221-1	Sequence 1, Appl
95	15	57.7	731	4	US-09-144-776B-11	Sequence 11, Appl
96	15	57.7	827	4	US-09-221-017B-265	Sequence 265, Ap
97	15	57.7	1050	4	US-09-543-236-753	Sequence 753, Ap
98	15	57.7	1154	4	US-09-489-039A-3477	Sequence 3477, Ap
99	15	57.7	1464	4	US-09-543-681A-2859	Sequence 2859, Ap
100	15	57.7	1464	4	US-09-543-681A-2859	Sequence 2859, Ap

us-09-877-819b-37.rni

Wed Apr 21 10:46:47 2004

Sequence 20, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 29, Appli

14.4 55.4 1035 4 US-08-778-717-20  
14.4 55.4 1044 3 US-09-248-528-6  
14.4 55.4 1044 3 US-09-549-108-6  
14.4 55.4 1044 3 US-09-549-111-6  
14.4 55.4 1044 3 US-09-549-106-6  
14.4 55.4 1044 3 US-09-550-394-6  
14.4 55.4 1252 4 US-09-561-756-29

# ALIGNMENTS

## RESULT 1

US-09-171-337A-2  
; Sequence 2, Application US/09171337A

; Patent No. 6300095

## GENERAL INFORMATION:

APPLICANT: BARREDO FUENTE, Jose Luis  
RODRIGUEZ SAIZ, Marta  
COLLADOS DE LA VIEJA, Alfonso J.  
MORENO VALLE, Migenl Angel  
SALTO MALDONADO, Francisco  
DIEZ GARCIA, Bruno

TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
AND -ACTIN AND THEIR USE IN FILAMENTOUS  
FUNGI EXPRESSION, SECRETION AND ANTISENSE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSER: LADAS & PARRY  
STREET: 26 WEST 61 STREET  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: 3-1/4" Disk 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11  
SOFTWARE: WordPerfect 8 for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/171,337A  
FILING DATE: 14-May-1999  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/ES98/00056  
FILING DATE: 5-MAR-1998  
APPLICATION NUMBER: ES9700482  
FILING DATE: 5-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: MASS, Clifford J.

REGISTRATION NUMBER: 30,086

(C) REF./DOCKET NO.: U-011948-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 233288

INFORMATION FOR SEQ ID NO: 2

SEQUENCE CHARACTERISTICS:

LENGTH: 5240 base pairs

TYPE: nucleotides

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Penicillium chrysogenum

IMMEDIATE SOURCE:

CLONE: plasmids pALP295 and pALP 388

FEATURE:

NAME/KEY: coding sequence

LOCATION: 1324 3111

FEATURE: OTHER INFORMATION: hex gene

Sequence 143, App  
Sequence 4, Appli  
Sequence 1, Appli  
Sequence 40, Appli  
Sequence 173, App  
Sequence 185, App  
Sequence 17, Appli  
Sequence 14, Appli  
Sequence 255, App  
Sequence 559, App  
Sequence 16287, A  
Sequence 16198, A  
Sequence 365, Ap  
Sequence 36, Appli  
Sequence 1, Appli  
Sequence 10, Appli  
Sequence 901, App  
Sequence 21, Appli  
Sequence 1335, Ap  
Sequence 25, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 27, Appli  
Sequence 252, App  
Sequence 69, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 79, Appli  
Sequence 137, App  
Sequence 137, App  
Sequence 137, App  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 10569, A  
Sequence 205, App  
Sequence 3058, Ap  
Sequence 3375, Ap  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 10288, A  
Sequence 1, Appli  
Sequence 14424, A  
Sequence 3311, Ap  
Sequence 528, App  
Sequence 3432, Ap  
Sequence 5, Appli  
Sequence 4904, Ap  
Sequence 2685, Ap  
Sequence 6, Appli  
Sequence 28, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 353, App  
Sequence 353, App  
Sequence 353, App  
Sequence 316, App  
Sequence 263, App  
Sequence 985, Ap  
Sequence 933, App  
Sequence 2543, Ap  
Sequence 3, Appli  
Sequence 9, Appli  
Sequence 2, Appli  
Sequence 1381, Ap  
Sequence 4095, Ap  
Sequence 1140, Ap

us-09-877-819b-37.rni

Wed Apr 21 10:46:47 2004

```
;
;
; FEATURE:
; OTHER INFORMATION: hex gene
; SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-631-022-2
Query Match 65.4%; Score 17; DB 4; Length 5240;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CAACTATGCGCGGTTGTACAGAC 26
| | | | | | | | | | | | | | | | | |
Db 2319 CAACTGTTCACGTTGTGCAGAC 2343

RESULT 3
US-09-631-022-2
; Sequence 320, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: DOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PS248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 320:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5897 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 320:
US-08-956-171E-320
Query Match 65.4%; Score 17; DB 4; Length 5897;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CACTTATGCGCGGTTGTACAGAC 26
| | | | | | | | | | | | | | | | | |
Db 2078 CATCTATGCGCGGTTGTGAACAC 2102

;
;
; FEATURE:
; OTHER INFORMATION: hex gene
; SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-171-337A-2
Query Match 65.4%; Score 17; DB 4; Length 5240;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CAACTATGCGCGGTTGTACAGAC 26
| | | | | | | | | | | | | | | | | |
Db 2319 CAACTGTTCACGTTGTGCAGAC 2343

RESULT 2
US-09-631-022-2
; Sequence 2, Application US/09631022
; Patent No. 6558921
; GENERAL INFORMATION:
; APPLICANT: BARRERO FUENTE, Jose Luis
; RODRIGUEZ SAIZ, Marta
; COLLADOS DE LA VIEJA, Alfonso J.
; MORENO VALLE, Migueal Angel
; SALTO MALDONADO, Francisco
; DIEZ GARCIA, Bruno
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE
; DEHYDROGENASE, -N-ACETYLHEXOSAMINIDASE
; AND -ACTIN AND THEIR USE IN FILAMENTOUS
; FUNGI EXPRESSION, SECRETION AND ANTISENSE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: LADAS & PARRY
; STREET: 26 WEST 61 STREET
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/4" Disk 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows for Workgroups 3.1-
; SOFTWARE: Wordperfect 8 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/631,022
; FILING DATE: 02-Aug-2000
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/171,337
; FILING DATE: 14-MAY-1999
; APPLICATION NUMBER: PCT/ES98/00056
; FILING DATE: 5-MAR-1998
; APPLICATION NUMBER: ES9700462
; FILING DATE: 5-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MASS, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NO.: U-02886-6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 233288
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5240 base pairs
; TYPE: nucleotides
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Penicillium chrysogenum
; IMMEDIATE SOURCE:
; CLONE: plasmids pALP295 and pALP 388
; FEATURE:
; NAME/KEY: coding sequence
; LOCATION: 1324 3111
```

Wed Apr 21 10:46:47 2004

us-09-877-819b-37.rni

```

RESULT 4
US-09-328-352-1668
; Sequence 1688, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1688
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1688

Query Match          64.6%; Score 16.8; DB 4; Length 1032;
Best Local Similarity 90.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CAATTATGCCGCTTCTTA 21
DB      320 CAATTATGCCGCTTGTGA 339

RESULT 5
US-08-674-351-1
; Sequence 1, Application US/08674351
; Patent No. 5831013
; GENERAL INFORMATION:
; APPLICANT: Bruenn, Jeremy A.
; APPLICANT: Yao, Wensheng
; TITLE OF INVENTION: CAPSID POLYPEPTIDES AND USE TO INHIBIT
; FILE OF INVENTION: VIRAL PACKAGING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,351
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19226/740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-674-351-1

Query Match          63.8%; Score 16.6; DB 2; Length 4580;
Best Local Similarity 82.6%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      4 ACTTATGCCGCTTGTACAGAC 26
DB      2318 ACTTATGCCGCTTGTACAGAC 2340

RESULT 6
US-09-702-705-1668
; Sequence 1668, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1668

Query Match          63.1%; Score 16.4; DB 4; Length 636;
Best Local Similarity 76.9%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 TCACCTATGCCGCTTGTACAGAC 26
DB      64 TCCGCTTATGCCGCTTGTGTGACAGAC 89

RESULT 7
US-09-736-457-1668
; Sequence 1668, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1668

Query Match          63.1%; Score 16.4; DB 4; Length 636;
Best Local Similarity 76.9%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```



```
QY 1 TCAACTTATGCGCTTCTTTGTGCAGAC 26
Db 64 TCCGCTTATGCGCTTCTTTGTGCAGAC 89

RESULT 8
US-09-614-124B-1668
; Sequence 1668, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Fanger, Gary
; APPLICANT: Lodes, Michael A.
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-1668

Query Match 63.1%; Score 16.4; DB 4; Length 636;
Best Local Similarity 76.9%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCTTCTTTGTGCAGAC 26
Db 64 TCCGCTTATGCGCTTCTTTGTGCAGAC 89

RESULT 9
US-09-671-325-1668
; Sequence 1668, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Fanger, Gary
; APPLICANT: Lodes, Michael A.
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1668

Query Match 63.1%; Score 16.4; DB 4; Length 636;
Best Local Similarity 76.9%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCTTCTTTGTGCAGAC 26
Db 64 TCCGCTTATGCGCTTCTTTGTGCAGAC 89

RESULT 10
US-09-702-705-316
; Sequence 316, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Fanger, Gary
; APPLICANT: Lodes, Michael A.
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 316
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-316

Query Match 63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCTTCTTTGTGCAGAC 26
Db 126 TCCGCTTATGCGCTTCTTTGTGCAGAC 151

RESULT 11
US-09-702-705-788
; Sequence 788, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Fanger, Gary
; APPLICANT: Lodes, Michael A.
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 788
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-788

Query Match 63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCTTCTTTGTGCAGAC 26
Db 126 TCCGCTTATGCGCTTCTTTGTGCAGAC 151
```

```

Query Match      63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Q1 1 TCAACTATGCGCGCTTTGTACAGAC 26
   ++ ++++++ ++++++ ++++++

```

Query Match	53.18;	Score 16.4;	DB 4;	Length 1633;
Best Local Similarity	76.94;	Pred. No. 57;		
Matches	20;	Conservative	0;	Mismatches 6;
				Indels 0;
Qy	1	TCAACTTATGCCGCGTTTGTACAGAC	26	
	126	TCGCGTTATGCGTTCTCTGTGCGAC	-51	

```
RESULT 16
US-09-671-325-316
; Sequence 316, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 316
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-316

Query Match          63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26
DB 126 TCCGCTTATGCCCTTCTTTGTGCAGAC 151

RESULT 17
US-09-671-325-788
; Sequence 788, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 788
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-788

Query Match          63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26
DB 126 TCCGCTTATGCCCTTCTTTGTGCAGAC 151

RESULT 18
US-09-671-325-788
; Sequence 788, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 788
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-788

Query Match          63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26
DB 126 TCCGCTTATGCCCTTCTTTGTGCAGAC 151
```

```
US-09-589-184-316
; Sequence 316, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 316
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-316

Query Match          63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26
DB 126 TCCGCTTATGCCCTTCTTTGTGCAGAC 151

RESULT 19
US-09-589-184-788
; Sequence 788, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 788
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-788

Query Match          63.1%; Score 16.4; DB 4; Length 1633;
Best Local Similarity 76.9%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGCTTTGTACAGAC 26
DB 126 TCCGCTTATGCCCTTCTTTGTGCAGAC 151

RESULT 20
US-09-596-002-26/c
; Sequence 26, Application US/09596002
; Patent No. 6632636
```

```

; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,172
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 34279
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inbyte template ID No. 6632636 26
; PUBLICATION INFORMATION:
; US-09-596-002-26

Query Match 63.1%; Score 16.4; DB 4; Length 34279;
Best Local Similarity 76.9%; Pred. No. 92;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCACCTTATGCCGCGTTGTACAGAC 26
Db 3476 TCAGGGTGTGACGGTTGTACAGAC 3451

RESULT 21
US-0734-674-3
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(202001)
; OTHER INFORMATION: n = A,T,C or G
; US-09-734-674-3

Query Match 63.1%; Score 16.4; DB 4; Length 202001;
Best Local Similarity 76.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCACCTTATGCCGCGTTGTACAGAC 26
Db 124154 TCACCTTATGATAGTTTATCCAGAC 124179

RESULT 22
US-09-328-352-3129
; Sequence 3129, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; US-09-328-352-3129
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; SEQ ID NO 3129
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-3129

Query Match 62.3%; Score 16.2; DB 4; Length 2571;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 CTTATGCCGCGTTGTACAGA 25
Db 1100 CTTATGCCGCGTTGTACAGA 1120

RESULT 23
US-08-916-421B-1/c
; Sequence 1, Application US/089164213
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)...(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)...(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)...(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)...(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)...(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)...(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)...(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)...(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)...(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)...(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)...(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:48:43 ; Search time 159.518 Seconds

Title: US-09-877-819B-37  
Perfect score: 26  
Sequence: 1 tcaacttatggcgggtttgtacagac 26

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	26	10	US-09-877-819B-37
2	26	100.0	254	10	US-09-877-819B-37
3	26	100.0	267	9	US-09-877-819B-37
4	26	100.0	267	9	US-09-877-819B-37
5	26	100.0	267	9	US-09-877-819B-37
6	26	100.0	267	15	US-10-040-862-143
7	26	100.0	267	15	US-10-040-862-143
8	26	100.0	267	15	US-10-040-862-143
9	26	100.0	267	16	US-10-057-475B-143
10	26	100.0	267	16	US-10-057-475B-143
11	26	100.0	267	16	US-10-057-475B-143
12	26	100.0	267	16	US-10-154-884B-143
13	26	100.0	267	16	US-10-154-884B-143
14	26	100.0	267	16	US-10-154-884B-143

26	100.0	272	9	US-09-795-692-42	Sequence 42, Appl
26	100.0	272	9	US-09-796-692-3554	Sequence 3554, Ap
26	100.0	272	15	US-10-040-862-42	Sequence 42, Appl
26	100.0	272	15	US-10-040-862-3554	Sequence 3554, Ap
26	100.0	272	16	US-10-057-475B-42	Sequence 42, Appl
26	100.0	272	16	US-10-057-475B-3554	Sequence 3554, Ap
26	100.0	272	16	US-10-154-884B-42	Sequence 42, Appl
26	100.0	272	16	US-10-154-884B-3554	Sequence 3554, Ap
26	100.0	272	16	US-09-796-692-580	Sequence 280, App
26	100.0	294	9	US-09-796-692-530	Sequence 3270, Ap
26	100.0	294	9	US-09-796-692-3270	Sequence 280, App
26	100.0	294	15	US-10-040-862-280	Sequence 280, App
26	100.0	294	15	US-10-040-862-530	Sequence 530, App
26	100.0	294	15	US-10-040-862-3270	Sequence 3270, Ap
26	100.0	294	16	US-10-057-475B-280	Sequence 280, App
26	100.0	294	16	US-10-057-475B-530	Sequence 530, App
26	100.0	294	16	US-10-057-475B-3270	Sequence 3270, Ap
26	100.0	294	16	US-10-154-884B-280	Sequence 280, App
26	100.0	294	16	US-10-154-884B-530	Sequence 530, App
26	100.0	294	16	US-10-154-884B-3270	Sequence 3270, Ap
26	100.0	410	10	US-09-918-995-36460	Sequence 37140, A
26	100.0	490	10	US-09-918-995-37140	Sequence 715, App
26	100.0	523	15	US-10-102-524-715	Sequence 88, Appl
26	100.0	601	15	US-10-084-827-88	Sequence 138, App
26	100.0	1259	13	US-10-220-120-138	Sequence 351, App
26	100.0	1348	9	US-09-925-302-351	Sequence 351, App
26	100.0	1348	13	US-09-925-302-351	Sequence 1043, Ap
26	100.0	14646	10	US-09-960-705-1043	Sequence 691, App
26	100.0	14646	10	US-09-873-319-691	Sequence 1, Appli
26	100.0	81.5	15	US-10-312-841-1	Sequence 2, Appli
26	100.0	3673778	15	US-10-312-841-2	Sequence 324, App
26	100.0	3673778	15	US-10-297-068-324	Sequence 59064, A
26	100.0	18	16	US-10-027-632-59064	Sequence 59064, A
26	100.0	18	16	US-10-027-632-59064	Sequence 7583, Ap
26	100.0	1601042	13	US-09-974-300-7583	Sequence 1121, Ap
26	100.0	1601042	13	US-09-237-183A-1119	Sequence 1119, Ap
26	100.0	197	10	US-09-237-183A-1119	Sequence 1120, Ap
26	100.0	296	10	US-09-237-183A-1120	Sequence 1134, Ap
26	100.0	307	10	US-09-237-183A-1134	Sequence 15638, A
26	100.0	420	10	US-10-425-114-15638	Sequence 85420, A
26	100.0	2115	13	US-10-424-599-85420	Sequence 115944, A
26	100.0	435	13	US-10-424-599-115944	Sequence 7077, Ap
26	100.0	797	13	US-10-425-114-7077	Sequence 32006, A
26	100.0	1512	13	US-10-425-114-32006	Sequence 5397, Ap
26	100.0	1783	13	US-10-425-114-5397	Sequence 34121, A
26	100.0	2032	13	US-10-282-122A-34121	Sequence 28843, A
26	100.0	3174	13	US-10-282-122A-28843	Sequence 320, App
26	100.0	4404	13	US-08-781-986A-320	Sequence 320, App
26	100.0	5897	8	US-10-329-624-320	Sequence 63, Appli
26	100.0	5897	13	US-09-861-451A-63	Sequence 2957, Ap
26	100.0	486	9	US-10-398-221-2957	Sequence 84681, A
26	100.0	1173	16	US-10-424-599-84681	Sequence 222578, A
26	100.0	536	13	US-10-027-632-222578	Sequence 110625, A
26	100.0	588	13	US-10-027-632-222578	Sequence 33968, A
26	100.0	588	16	US-10-424-599-110625	Sequence 25267, A
26	100.0	691	13	US-10-369-493-33968	Sequence 7938, Ap
26	100.0	1359	16	US-10-282-122A-25267	Sequence 84682, A
26	100.0	1539	13	US-10-425-114-7938	Sequence 375, App
26	100.0	2614	13	US-10-424-599-84682	Sequence 16333, A
26	100.0	3275	13	US-09-070-927A-375	Sequence 8759, A
26	100.0	7664	15	US-10-029-386-16333	Sequence 386, App
26	100.0	224	15	US-10-424-599-81759	Sequence 183789, A
26	100.0	267	13	US-10-116-712-386	Sequence 183789, A
26	100.0	398	15	US-10-027-632-183789	Sequence 1632, Ap
26	100.0	412	13	US-10-027-632-183789	Sequence 2633, Ap
26	100.0	412	16	US-10-102-524-1632	Sequence 108668, A
26	100.0	432	15	US-10-029-386-2633	Sequence 108668, A
26	100.0	551	15	US-10-027-632-108668	Sequence 72467, A
26	100.0	594	13	US-10-027-632-108668	Sequence 299113, A
26	100.0	594	16	US-10-027-632-72467	Sequence 72467, A
26	100.0	596	13	US-10-027-632-299113	Sequence 299113, A
26	100.0	596	13	US-10-027-632-72467	Sequence 299113, A
26	100.0	596	16	US-10-027-632-72467	Sequence 299113, A
26	100.0	63.1	16	US-10-027-632-299113	Sequence 299113, A

```

88 16.4 63.1 624 9 US-09-902-941-1914 Sequence 1914, Ap
89 16.4 63.1 624 9 US-09-849-626-1914 Sequence 1914, Ap
90 16.4 63.1 624 13 US-10-283-017-1914 Sequence 1914, Ap
91 16.4 63.1 624 15 US-10-017-754-1914 Sequence 1914, Ap
92 16.4 63.1 624 15 US-10-113-872-1914 Sequence 1914, Ap
93 16.4 63.1 636 9 US-09-736-457-1668 Sequence 1668, Ap
94 16.4 63.1 636 9 US-09-902-941-1668 Sequence 1668, Ap
95 16.4 63.1 636 9 US-09-849-626-1668 Sequence 1668, Ap
96 16.4 63.1 636 13 US-10-283-017-1668 Sequence 1668, Ap
97 16.4 63.1 636 15 US-10-017-754-1668 Sequence 1668, Ap
98 16.4 63.1 636 15 US-10-113-872-1668 Sequence 1668, Ap
99 16.4 63.1 714 13 US-10-424-599-69824 Sequence 69824, A
100 16.4 63.1 1064 13 US-10-424-599-12667 Sequence 12667, A
101 16.4 63.1 1126 13 US-10-428-114-2114 Sequence 2114, Ap
102 16.4 63.1 1166 13 US-10-428-114-35400 Sequence 35400, A
103 16.4 63.1 1260 16 US-10-369-493-41879 Sequence 41879, A
104 16.4 63.1 1619 9 US-09-070-927A-563 Sequence 563, Ap
105 16.4 63.1 1633 9 US-09-736-457-316 Sequence 316, Ap
106 16.4 63.1 1633 9 US-09-736-457-788 Sequence 788, Ap
107 16.4 63.1 1633 9 US-09-502-941-316 Sequence 316, Ap
108 16.4 63.1 1633 9 US-09-502-941-788 Sequence 788, Ap
109 16.4 63.1 1633 9 US-09-849-626-316 Sequence 316, Ap
110 16.4 63.1 1633 9 US-09-849-626-788 Sequence 788, Ap
111 16.4 63.1 1633 10 US-09-476-300-316 Sequence 316, Ap
112 16.4 63.1 1633 13 US-10-342-887-1143 Sequence 1143, Ap
113 16.4 63.1 1633 13 US-10-172-118-1143 Sequence 116, Ap
114 16.4 63.1 1633 13 US-10-283-017-316 Sequence 316, Ap
115 16.4 63.1 1633 13 US-10-283-017-788 Sequence 788, Ap
116 16.4 63.1 1633 15 US-10-017-754-316 Sequence 316, Ap
117 16.4 63.1 1633 15 US-10-017-754-788 Sequence 788, Ap
118 16.4 63.1 1633 15 US-10-113-872-316 Sequence 316, Ap
119 16.4 63.1 1633 15 US-10-113-872-788 Sequence 788, Ap
120 16.4 63.1 3159 16 US-10-369-493-28565 Sequence 28565, A
121 16.4 63.1 7880 9 US-09-764-860-960 Sequence 960, Ap
122 16.4 63.1 7880 9 US-09-764-860-961 Sequence 961, Ap
123 16.4 63.1 7880 15 US-10-074-095-960 Sequence 960, Ap
124 16.4 63.1 7880 15 US-10-074-095-961 Sequence 960, Ap
125 16.4 63.1 7880 16 US-10-212-872-960 Sequence 961, Ap
126 16.4 63.1 7880 16 US-10-212-872-961 Sequence 961, Ap
127 16.4 63.1 21636 9 US-09-416-384A-3 Sequence 3, Appli
128 16.4 63.1 34279 13 US-10-672-787-26 Sequence 26, Appl
129 16.4 63.1 63248 13 US-10-087-192-1879 Sequence 1879, Ap
130 16.4 63.1 88853 9 US-10-087-192-85 Sequence 85, Appli
131 16.4 63.1 202001 9 US-09-734-674-3 Sequence 3, Appli
132 16.4 63.1 202001 15 US-10-274-590-3 Sequence 3, Appli
133 16.2 62.3 25 15 US-10-098-263B-62452 Sequence 62452, A
134 16.2 62.3 431 13 US-10-424-599-129450 Sequence 129450, A
135 16.2 62.3 507 15 US-10-252-157-275 Sequence 275, Appl
136 16.2 62.3 591 15 US-10-029-386-150 Sequence 150, Appl
137 16.2 62.3 889 13 US-10-424-599-53165 Sequence 53165, A
138 16.2 62.3 1842 9 US-09-815-242-6128 Sequence 6128, Ap
139 16.2 62.3 1842 13 US-10-282-122A-20434 Sequence 20434, A
140 16.2 62.3 2022 13 US-10-282-122A-41966 Sequence 41966, A
141 16.2 62.3 2396 13 US-10-282-122A-8502 Sequence 8502, Ap
142 16.2 62.3 3177 13 US-10-282-122A-36107 Sequence 36107, A
143 16.2 62.3 272 13 US-10-424-599-106911 Sequence 106911, A
144 16.1 61.5 313 9 US-09-764-847-1180 Sequence 1180, Ap
145 16.1 61.5 585 13 US-10-092-154-1180 Sequence 1180, Ap
146 16.1 61.5 585 13 US-10-027-632-320083 Sequence 320083, A
147 16.1 61.5 585 16 US-10-027-632-320083 Sequence 320083, A
148 16.1 61.5 591 13 US-10-027-632-35907 Sequence 35907, A
149 16.1 61.5 591 13 US-10-027-632-69244 Sequence 69244, A
150 16.1 61.5 591 16 US-10-027-632-69244 Sequence 69244, A
151 16.1 61.5 591 16 US-10-027-632-69244 Sequence 69244, A
152 16.1 61.5 612 13 US-10-424-599-130762 Sequence 130762, A
153 16.1 61.5 732 13 US-10-282-122A-6482 Sequence 6482, Ap
154 16.1 61.5 755 16 US-10-287-274-220 Sequence 220, Ap
155 16.1 61.5 1478 13 US-10-369-493-45971 Sequence 45971, A
156 16.1 61.5 1766 13 US-10-424-599-13195 Sequence 13195, A
157 16.1 61.5 2000 9 US-09-938-842A-4705 Sequence 4705, Ap
158 16.1 61.5 2100 11 US-10-113-262-3 Sequence 3, Appli
159 16.1 61.5 2100 14 US-10-113-262-3 Sequence 3, Appli

```

Wed Apr 21 10:46:48 2004

us-09-877-819b-37.rnpb

```

; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 128
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-328

Query Match 100.0%; Score 26; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 140 TCAACTTATGCGCGTTGTACAGAC 165

RESULT 5
US-09-796-692-7075
; Sequence 7075, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-143

Query Match 100.0%; Score 26; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 140 TCAACTTATGCGCGTTGTACAGAC 165

RESULT 4
US-09-796-692-328
; Sequence 328, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
```

us-09-877-819b-37.rnpb

Wed Apr 21 10:46:48 2004

```

US-10-040-862-143
Query Match      100.0%; Score 26; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTGTACAGAC 26
   |||||
DB 140 TCAACTTATGCGCGTTGTGTACAGAC 165

RESULT 7
US-10-040-862-328
; Sequence 328, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 328
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-328
Query Match      100.0%; Score 26; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTGTACAGAC 26
   |||||
DB 140 TCAACTTATGCGCGTTGTGTACAGAC 165

RESULT 8
US-10-040-862-7075
; Sequence 7075, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7075
Query Match      100.0%; Score 26; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTGTACAGAC 26
   |||||
DB 140 TCAACTTATGCGCGTTGTGTACAGAC 165

RESULT 6
US-10-040-862-143
; Sequence 143, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7075

```



```

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,962
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/220,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7075
LENGTH: 267
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-7075
Query Match 100.0%; Score 26; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATCGCGGTTGTGTACAGAC 26
Db 140 TCAACTTATCGCGGTTGTGTACAGAC 165

RESULT 9
US-10-057-475B-143
Sequence 143, Application US/10057475B
Publication No. US20040002068A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
CURRENT FILING DATE: 2002-01-22
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22

```



Query Match 100.0%; Score 26; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.0044; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 26; DB 16; Length 267;

Best Local Similarity 100.0%; Pred. No. 0.0044; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 26; DB 16; Length 267;

Best Local Similarity 100.0%; Pred. No. 0.0044; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 26; DB 16; Length 267;

Best Local Similarity 100.0%; Pred. No. 0.0044; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 16
US-09-796-692-3554
; Sequence 3554, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-42

Query Match 100.0%; Score 26; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTGTACAGAC 26
Db 145 TCAACTTATGCGCGTTGTGTACAGAC 170

RESULT 17
US-10-040-862-42
; Sequence 42, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-3554

Query Match 100.0%; Score 26; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTGTACAGAC 26
Db 145 TCAACTTATGCGCGTTGTGTACAGAC 170

RESULT 18
US-10-040-862-3554
; Sequence 3554, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
```

RESULT 21  
US-10-154-884B-42  
; Sequence 42, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane

```
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-42

Query Match 100.0%; Score 26; DB 15; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 145 TCAACTTATGCGCGTTGTACAGAC 170

RESULT 22
US-10-154-884B-3554
; Sequence 3554, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,126
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-42
```

```
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3554
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-3554

Query Match 100.0%; Score 26; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 145 TCAACTTATGCGCGTTGTACAGAC 170

RESULT 23
US-09-796-692-280/c
; Sequence 280, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 280
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-280

Query Match 100.0%; Score 26; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

1  TCAACTTATCCCGCGTTTGTACAGAC 26
128 TCAACTTATCCCGCGTTTGTACAGAC 103
DBB

RESULT 24
US-09-796-692-530/c
; Sequence 530, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mahnon, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.00i200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999

```

, PRIOR APPLICATION NUMBER: 60/202,084  
 , PRIOR FILING DATE: 2000-05-04  
 , PRIOR APPLICATION NUMBER: 60/206,201  
 , PRIOR FILING DATE: 2000-05-22  
 , PRIOR APPLICATION NUMBER: 60/218,950  
 , PRIOR FILING DATE: 2000-07-14  
 , PRIOR APPLICATION NUMBER: 60/222,903

```

; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-530

      Query Match      100.0%; Score 26; DB 9; Length 294;
      Best Local Similarity 100.0%; Pred. No. 0.0045;
      Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      TCAACTTATGCCGCGTTTGTACAGAC 26
          |||||
Db      128    TCAACTTATGCCGCGTTTGTACAGAC 103

RESULT 25
US-09-796-692-3270/c
; Sequence 3270, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692

```





```

; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US 60/190,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-057-475B-530

Query Match      100.0%; Score 26; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TCAACTTATGCGCGTTTGTACAGAC 26
      |||||
Db      128 TCAACTTATGCGCGTTTGTACAGAC 103

RESULT 31
US-10-057-475B-3270/C
; Sequence 3270, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950

US-10-057-475B-530/C
; Sequence 530, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
```

;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; Remaining prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 10979  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3270  
;; LENGTH: 294  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-057-475B-3270

Query Match 100.0%; Score 26; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
Db 128 TCAACTTATGCGCGTTGTACAGAC 103

RESULT 32  
US-10-154-884B-280/c  
; Sequence 280, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 280  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-280

Query Match 100.0%; Score 26; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
Db 128 TCAACTTATGCGCGTTGTACAGAC 103

RESULT 33  
US-10-154-884B-530/c  
; Sequence 530, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 530  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-154-884B-530

Query Match 100.0%; Score 26; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
Db 128 TCAACTTATGCGCGTTGTACAGAC 103

RESULT 34  
US-10-154-884B-3270/c  
; Sequence 3270, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17

```

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37140
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-37140

Query Match          100.0%; Score 26; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCAACTTATGCCGCGTTTGTCACGAC 26
        |||||||
DB      236 TCAACTTATGCCGCGTTTGTCACGAC 261

RESULT 37
US-10-102-524-715
; Sequence 715, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 715
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-715

Query Match          100.0%; Score 26; DB 15; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCAACTTATGCCGCGTTTGTCACGAC 26
        |||||||
DB      151 TCAACTTATGCCGCGTTTGTCACGAC 176

RESULT 38
US-10-084-817-88
; Sequence 88, Application US/10084817
; Publication No. US20030115009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365

```

SOFTWARE: PERL Program  
SEQ ID NO 88  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030119009A1 1090035.1  
US-10-084-817-86

Query Match 100.0%; Score 26; DB 15; Length 601;  
Best Local Similarity 100.0%; Pred. No. 0.0053;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
Db 184 TCAACTTATGCGCGTTGTACAGAC 209

## RESULT 39

US-10-220-120-138  
Sequence 138, Application US/10320120

Publication No. US20040048253A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: PANZER, Scott R.

APPLICANT: SPIRO, Peter A.

APPLICANT: BANVILLE, Steven C.

APPLICANT: SHAH, Purvi

APPLICANT: CHALUP, Michael S.

APPLICANT: CHEN, Simon C.

APPLICANT: CHEN, Alice

APPLICANT: D'SA, Steven A.

APPLICANT: AMSHEY, Stefan

APPLICANT: DAHL, Christopher R.

APPLICANT: DAM, Tam C.

APPLICANT: DANIELS, Susan E.

APPLICANT: DUFOUR, Gerard E.

APPLICANT: FLORES, Vincent

APPLICANT: FONG, Willy T.

APPLICANT: GREENAWALT, Lila B.

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: JONES, Anissa L.

APPLICANT: LIU, Tommy F.

APPLICANT: ROSEBERRY, Ann M.

APPLICANT: ROSK, Bruce H.

APPLICANT: RUSSO, Frank D.

APPLICANT: STOCKREHER, Theresa K.

APPLICANT: DAFFO, Abel

APPLICANT: WRIGHT, Rachel J.

APPLICANT: YAP, Pierre E.

APPLICANT: YU, Jimmy Y.

APPLICANT: BRADLEY, Diana L.

APPLICANT: BRATCHER, Shawn R.

APPLICANT: CHEN, Wensheng

APPLICANT: COHEN, Howard J.

APPLICANT: HODGSON, David M.

APPLICANT: LINCOLN, Stephen E.

APPLICANT: JACKSON, Stuart

TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: PT-1113 PCT

CURRENT APPLICATION NUMBER: US/10/220,120

PRIOR FILING DATE: 2002-08-26

PRIOR APPLICATION NUMBER: 60/184,777; 60/184,698; 60/184,774;

60/184,693; 60/184,771; 60/184,813; 60/184,776;

60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;

60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;

60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;

60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;

60/205,324; 60/205,286

PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;

2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;

2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;

2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;  
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;  
2000-05-17; 2000-05-17  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: PERL Program  
SEQ ID NO 138  
LENGTH: 1259  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:1169865.1:2000MAY01  
US-10-220-120-138

Query Match 100.0%; Score 26; DB 13; Length 1259;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGAC 26  
Db 326 TCAACTTATGCGCGTTGTACAGAC 351

## RESULT 40

US-09-925-302-351

Sequence 351, Application US/09925302

Patent No. US20020044941A1

GENERAL INFORMATION:

APPLICANT: ROSEN ET AL.

FILE REFERENCE: PA104

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

CURRENT APPLICATION NUMBER: US/09/925,302

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US90/05918

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 896

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 351

LENGTH: 1348

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (12)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (24)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (36)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1294)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1329)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1307)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (1318)

OTHER INFORMATION: n equals a,t,g, or c

```
Db 264 TCAACTTATGCCGCGTTTGTACAGAC 289
|||||
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-960-706-1043

Query Match 100.0%; Score 26; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCCGCGTTTGTACAGAC 26
|||||
Db 6345 TCAACTTATGCCGCGTTTGTACAGAC 6370

RESULT 43
US-09-873-319-691
; Sequence 691, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 691
; LENGTH: 14646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X03100
; NAME/KEY: unsure
; LOCATION: (1)..(14646)
; OTHER INFORMATION: n = a or c or g or t
US-09-873-319-691

Query Match 100.0%; Score 26; DB 10; Length 14646;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCCGCGTTTGTACAGAC 26
|||||
Db 6345 TCAACTTATGCCGCGTTTGTACAGAC 6370

RESULT 44
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
```

```
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
```

```
Query Match      81.5%; Score 21.2; DB 15; Length 3673778;
Best Local Similarity 88.5%; Pred. No. 9.3;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 3298050 TCAACTTATGCGCGTTGTACAGAC 3298025
```

```
RESULT 45
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2
```

```
Query Match      75.4%; Score 19.5; DB 15; Length 3673778;
Best Local Similarity 84.6%; Pred. No. 59;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 375729 TTAATTTATGCGGTTGTATAGAC 375754
```

```
RESULT 46
US-10-297-068-324
; Sequence 324, Application US/10297068
; Publication No. US20030228585A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: KAGIYA, Taeko
; APPLICANT: ICHIHARA, Tatsuo
; APPLICANT: Matsumura, Yoshiyuki
; APPLICANT: MORIYA, Shogo
; APPLICANT: NISHIDA, Michio
; TITLE OF INVENTION: KIT AND METHOD FOR DETERMINING HLA TYPES
; FILE REFERENCE: 13140P1174
; CURRENT APPLICATION NUMBER: US/10/297,068
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: JP 2000-164798
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 1298
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 324
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:capture
US-10-297-068-324
```

```
Query Match      69.2%; Score 18; DB 16; Length 18;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 5 CTTATGCGCGTTGTATC 22
Db 1 CTTATGCGCGTTGTATC 18
```

```
RESULT 47
US-10-027-632-59064
; Sequence 59064, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59064
; LENGTH: 1601042
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1601042)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064
```

```
Query Match      58.5%; Score 17.8; DB 13; Length 1601042;
Best Local Similarity 76.0%; Pred. No. 3.9e+02;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 TCAACTTATGCGCGTTGTACAGA 25
Db 1453122 WCRGCTTATGCGGCTGTACTACAGA 1453146
```

```
RESULT 48
US-10-027-632-59064
; Sequence 59064, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59064
; LENGTH: 1601042
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(1601042)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

```

```

Query Match      68.5%; Score 17.8; DB 16; Length 1601042;
Best Local Similarity 76.0%; Pred. No. 3.9e+02;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTACAGA 25
Db 1453122 WCAGCTTATGCGCGTGTATCAGA 1453146

```

```

RESULT 49
US-09-974-300-7583
; Sequence 7583, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7583
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-7583

```

```

Query Match      66.9%; Score 17.4; DB 9; Length 534;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 ATGCGCGTTGTACAGAC 26
Db 11 ATGCGCGTTGTACTGAC 29

```

```

RESULT 50
US-09-237-183A-1121/c
; Sequence 1121, Application US/09237183A
; Publication No. US20030135870A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, No. US20030135870A1-dine
; APPLICANT: Fisher, Dane K.
; APPLICANT: Liu, Jingdong

```

```

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; FILE REFERENCE: 38-21(15089)B
; CURRENT APPLICATION NUMBER: US/09/237,183A
; CURRENT FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 2814
; SEQ ID NO 1121
; LENGTH: 197
; TYPE: DNA
; ORGANISM: Zea mays
US-09-237-183A-1121

```

```

Query Match      66.2%; Score 17.2; DB 10; Length 197;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTAC 22
Db 54 TCAACATATGCTGCGTTGTGC 33

```

```

RESULT 51
US-09-237-183A-1119/c
; Sequence 1119, Application US/09237183A
; Publication No. US20030135870A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, No. US20030135870A1-dine
; APPLICANT: Fisher, Dane K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; FILE REFERENCE: 38-21(15089)B
; CURRENT APPLICATION NUMBER: US/09/237,183A
; CURRENT FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 2814
; SEQ ID NO 1119
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Zea mays
US-09-237-183A-1119

```

```

Query Match      66.2%; Score 17.2; DB 10; Length 296;
Best Local Similarity 86.4%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGTTGTAC 22
Db 213 TCAACATATGCTGCGTTGTGC 192

```

```

RESULT 52
US-09-237-183A-1120/c
; Sequence 1120, Application US/09237183A
; Publication No. US20030135870A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, No. US20030135870A1-dine
; APPLICANT: Fisher, Dane K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; FILE REFERENCE: 38-21(15089)B
; CURRENT APPLICATION NUMBER: US/09/237,183A
; CURRENT FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 2814
; SEQ ID NO 1120
; LENGTH: 307
; TYPE: DNA

```

```
/ ORGANISM: Zea mays
US-09-237-183A-1120

Query Match
Best Local Similarity 66.2%; Score 17.2; DB 10; Length 307;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGTTGTAC 22
Db 57 TCAACATATGCTGCGTTGTGC 36

RESULT 53
US-09-237-183A-1134/c
; Sequence 1134, Application US/09237183A
; Publication No. US20030135870A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, No. US20030135870A1d:ne
; APPLICANT: Fisher, Dane K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; TITLE OF INVENTION: Sucrose Pathway
; FILE REFERENCE: 38-21(15089)B
; CURRENT APPLICATION NUMBER: US/09/237,183A
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 2814
; SEQ ID NO 1134
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Zea mays
US-09-237-183A-1134

Query Match
Best Local Similarity 66.2%; Score 17.2; DB 10; Length 420;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGTTGTAC 22
Db 236 TCAACATATGCTGCGTTGTGC 215

RESULT 54
US-10-425-114-15638/c
; Sequence 15638, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15638
; LENGTH: 2115
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LTB3060-044-G7_FLI
US-10-425-114-15638

Query Match
Best Local Similarity 66.2%; Score 17.2; DB 13; Length 2115;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGTTGTAC 22
```

```
Db 1127 TCAACATATGCTGCGTTGTGC 1106

RESULT 55
US-10-424-599-85420/c
; Sequence 85420, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85420
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48-49C.1
US-10-424-599-85420

Query Match
Best Local Similarity 65.4%; Score 17; DB 13; Length 435;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGTTGTACAG 25
Db 59 TAAATTATGCCACGTTTGACAAA 35

RESULT 56
US-10-424-599-115944
; Sequence 115944, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 115944
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75708C.1
US-10-424-599-115944

Query Match
Best Local Similarity 65.4%; Score 17; DB 13; Length 797;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CAACATATGCGCGTTGTACAG 26
Db 449 CAACATATGCTGCTGTACAGCC 473

RESULT 57
US-10-425-114-7077/c
; Sequence 7077, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
```





Query Match 65.4%; Score 17; DB 13; Length 3174;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CAACTATGCCCGGTTGTACAGAC 26  
DB 1022 CATCTATGCAGCGTTGAACCAAC 1046

Search completed: April 20, 2004, 12:56:01  
Job time : 176.518 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 07:16:28 ; Search time 1654.67 Seconds

(without alignments)  
469.227 Million cell updates/sec

Title: US-09-877-819b-37

Perfect score: 26

Sequence: 1 tcaacttatgcgcgtttgtacagac 26

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfum:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pla:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	159	4	T91602
2	26	100.0	229	10	BE140710
3	26	100.0	256	10	BE159476
4	26	100.0	265	10	BE242285

26	100.0	289	9	AA360953
28	100.0	290	14	CD706205
26	100.0	295	14	CD694567
26	100.0	335	12	BM772902
26	100.0	338	9	AA3323639
26	100.0	366	14	T94759
26	100.0	387	14	CB267061
26	100.0	405	9	AA244273
26	100.0	415	9	AA838010
26	100.0	422	13	CO3540
26	100.0	423	14	CD102141
26	100.0	427	9	AW085969
26	100.0	437	12	BM767805
26	100.0	469	12	BM766631
26	100.0	484	10	BF819626
26	100.0	491	12	BM694247
26	100.0	494	14	CD698716
26	100.0	511	14	CD707609
26	100.0	514	14	CD698819
26	100.0	515	12	BM769742
26	100.0	519	12	BG541135
26	100.0	526	14	CD705087
26	100.0	533	14	CD687511
26	100.0	534	12	BG756165
26	100.0	540	14	CD695435
26	100.0	544	10	BR874055
26	100.0	545	13	BQ267707
26	100.0	545	13	BU783326
26	100.0	546	10	AW351777
26	100.0	547	10	AW406086
26	100.0	547	14	CD708950
26	100.0	557	12	BG535978
26	100.0	557	14	CD687141
26	100.0	563	14	CB265399
26	100.0	571	14	CA942442
26	100.0	579	12	BM831052
26	100.0	582	12	BM737984
26	100.0	585	9	AV706521
26	100.0	585	13	BU783392
26	100.0	588	12	BI911442
26	100.0	591	14	CB551184
26	100.0	592	12	BM876262
26	100.0	594	12	BQ26106
26	100.0	595	14	CB693703
26	100.0	597	14	CD684394
26	100.0	598	14	CD695219
26	100.0	599	13	BQ270752
26	100.0	601	10	BF725429
26	100.0	602	13	BU783579
26	100.0	602	14	CD699993
26	100.0	610	14	CA405960
26	100.0	626	12	BG545380
26	100.0	627	12	BG715375
26	100.0	627	14	CB267128
26	100.0	629	14	CD691540
26	100.0	632	14	CB550782
26	100.0	635	10	BE712825
26	100.0	635	12	BG709234
26	100.0	643	12	BM767502
26	100.0	645	12	BI769913
26	100.0	645	14	CB551104
26	100.0	658	9	AV763677
26	100.0	659	14	CD687077
26	100.0	663	14	CD701461
26	100.0	667	9	AL048565
26	100.0	669	13	BU071266
26	100.0	672	14	CD692593
26	100.0	676	12	BG570300
26	100.0	681	14	CB554193
26	100.0	683	9	AV759427
26	100.0	684	14	CD693272
26	100.0	685	10	B3973719
26	100.0	693	12	BG685335

AA360953	EST70157
CD706205	EST22732
CD694567	EST11090
BM772902	K-EST0057
AA3323639	EST26729
T94759	ve37fl2.r1
CB267061	1005567.H
AA244273	nc06d06.r
AA838010	o689d07.s
CO3540	CO3540 Huma
CD102141	AGENCOURT
AW085969	xc76d02.x
BM767805	K-EST0050
BM766631	K-EST0048
BF819626	MRI-RT002
BM694247	UI-E-C11-
CD698716	EST15239
CD707609	EST24136
CD698819	EST15342
BM769742	K-EST0053
BG541135	602569911
CD705087	EST21614
CD687511	EST4032.h
BG756165	602713466
CD695435	EST11958
BE874055	601484433
BU783326	in03f04.Y
AW351777	RC3-CT019
AW406086	UI-HF-BL0
CD708950	EST23477
BG535978	602564150
CD687141	EST3662.h
CB265399	1004304.H
CA942442	1r58a06.Y
BM831052	K-EST0104
BM737984	K-EST0002
AV706521	AV706521
BU783392	it03f10.Y
BI911442	603063356
CB551184	MMSPO068
BM876262	ij60c02.X
BQ26106	HNC50-1.B
CD693703	EST10226
CD684394	EST314.hu
CD695219	EST11742
BQ270752	ik04d01.Y
BF725429	bx15h06.Y
BU783579	in06c10.Y
CD699993	EST16517
CA405960	1002109.H
BG545380	602572611
BG715375	602675513
CB267128	1006034.H
CD691540	EST8063.h
CB550782	MMSPO018
BE712825	MRO-HT069
BG709234	602674620
BM767502	K-EST0049
BI769913	603060181
CB551104	MMSPO061
AV763677	AV763677
CD687077	EST3598.h
CD701461	EST17885
AL048565	DKF25586P
BU071266	im41b05.X
CD692593	EST9122.h
BG570300	602590723
CB554193	MMSPO008
AV759427	AV759427
CD693272	EST9795.h
B3973719	601680707
BG685335	604637767

78 26 100.0 704 9 AV707666  
 79 26 100.0 706 12 BG431973  
 80 26 100.0 708 12 BG539584  
 81 26 100.0 709 12 BG758164  
 82 26 100.0 724 10 BE439689  
 83 26 100.0 732 12 BG699921  
 84 26 100.0 732 14 CB997879  
 85 26 100.0 734 9 AV733676  
 86 26 100.0 735 12 BI762063  
 87 26 100.0 735 14 CB551496  
 88 26 100.0 743 14 CB956955  
 89 26 100.0 743 14 CD104670  
 90 26 100.0 744 12 BI597401  
 91 26 100.0 745 14 CB984378  
 92 26 100.0 751 12 BG759996  
 93 26 100.0 758 14 CB985816  
 94 26 100.0 771 12 BI753761  
 95 26 100.0 776 12 BI227278  
 96 26 100.0 785 9 AV762204  
 97 26 100.0 789 12 BG538956  
 98 26 100.0 791 13 BU596174  
 99 26 100.0 796 14 CB997069  
 100 26 100.0 797 12 BG758487  
 101 26 100.0 802 12 BG757550  
 102 26 100.0 806 14 CB993192  
 103 26 100.0 809 14 CB993270  
 104 26 100.0 810 9 AV764215  
 105 26 100.0 811 12 BI766898  
 106 26 100.0 817 12 BG758419  
 107 26 100.0 821 12 BI668794  
 108 26 100.0 822 14 BG341032  
 109 26 100.0 823 14 CB991431  
 110 26 100.0 826 12 BG431923  
 111 26 100.0 830 14 CB994002  
 112 26 100.0 832 14 CB958376  
 113 26 100.0 837 12 BI517599  
 114 26 100.0 843 10 BF975896  
 115 26 100.0 843 12 BI261592  
 116 26 100.0 851 14 CD558780  
 117 26 100.0 854 12 BG756171  
 118 26 100.0 857 13 BG890384  
 119 26 100.0 859 12 BG546165  
 120 26 100.0 860 12 BG755507  
 121 26 100.0 864 10 BF663966  
 122 26 100.0 871 12 BI767351  
 123 26 100.0 872 13 BQ721555  
 124 26 100.0 873 12 BG540219  
 125 26 100.0 878 14 CB986657  
 126 26 100.0 879 13 BU529093  
 127 26 100.0 880 9 AU139061  
 128 26 100.0 881 14 CD558622  
 129 26 100.0 884 12 BI545349  
 130 26 100.0 889 14 CD516536  
 131 26 100.0 901 13 BQ955105  
 132 26 100.0 915 10 BF974675  
 133 26 100.0 917 12 BG975894  
 134 26 100.0 923 13 BQ642494  
 135 26 100.0 926 13 BQ642196  
 136 26 100.0 930 13 BQ712102  
 137 26 100.0 943 13 BQ060004  
 138 26 100.0 959 10 BF976257  
 139 26 100.0 959 13 BQ057748  
 140 26 100.0 960 13 BQ894372  
 141 26 100.0 964 13 BQ642742  
 142 26 100.0 966 13 BQ706261  
 143 26 100.0 967 12 BI553587  
 144 26 100.0 980 10 BF974318  
 145 26 100.0 980 13 BQ056978  
 146 26 100.0 990 12 BG686860  
 147 26 100.0 998 13 BU899275  
 148 26 100.0 1005 13 BQ062257  
 149 26 100.0 1007 13 BQ057757  
 150 26 100.0 1015 13 BQ060930

151 26 100.0 1025 12 BQ054502  
 152 26 100.0 1026 13 BQ061344  
 153 26 100.0 1027 13 BQ064032  
 154 26 100.0 1030 13 BQ063085  
 155 26 100.0 1043 13 BQ882939  
 156 26 100.0 1055 14 CD517144  
 157 26 100.0 1061 12 BQ054682  
 158 26 100.0 1083 13 BU594971  
 159 26 100.0 1108 13 BQ072806  
 160 26 100.0 1111 14 CD517173  
 161 26 100.0 1148 10 BF974697  
 162 26 100.0 1201 13 BX363746  
 163 26 100.0 1217 12 BM544066  
 164 26 100.0 1722 13 BQ057412  
 165 24.4 93.8 168 14 CD686854  
 166 24.4 93.8 249 14 CD707370  
 167 24.4 93.8 390 14 CD709023  
 168 24.4 93.8 464 14 CD705070  
 169 24.4 93.8 559 14 CD700948  
 170 24.4 93.8 618 14 CD686880  
 171 24.4 93.8 623 14 CD767511  
 172 24.4 93.8 735 12 BG484403  
 173 24.4 93.8 786 12 BG427527  
 174 24.4 93.8 790 12 BI837330  
 175 23 88.5 230 13 C03623  
 176 22.8 87.7 177 10 BG025808  
 177 22.8 87.7 318 10 AW404262  
 178 22.8 87.7 334 14 CD704858  
 179 22.8 87.7 408 12 BM834672  
 180 22.9 87.7 439 10 AW406315

## ALIGNMENTS

## RESULT 1

T91602

LOCUS

DEFINITION

T91602 159 bp mRNA linear EST 22-MAR-1995  
 Ye21b03.r1 Stratagene lung (#937210) Homo sapiens cDNA clone  
 IMAGE:118349 5' similar to gb:K01506 HLA CLASS II  
 HISTOCOMPATIBILITY ANTIGEN, DP(1) ALPHA CHAIN (HUMAN); mRNA  
 sequence.

ACCESSION T91602.1 GI:723515

VERSION T91602

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 159)  
 Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
 Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,  
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,X., Le,N.,  
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,P., Thierry-Mieg,J.,  
 Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.  
 and Marra,M.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 97044478  
 8889549  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estowatson.wustl.edu  
 High quality sequence stops: 1:4  
 Source: IMAGE Consortium, LIND  
 This clone is available royalty-free through LIND; contact the  
 IMAGE Consortium (info@image.lind.gov) for further information.  
 Seq primer: M13RPI  
 High quality sequence stop: 114.

## FEATURES

source Location/Qualifiers  
 1. .159  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:486638"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:118349"  
 /sex="male"  
 /dev\_stage="72 years"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Stratagene lung (#937210)"  
 /notes="Organ: Lung; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT, normal lung. Average insert size: 1.3 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCGATTTTITTTTITTT 3'"

## ORIGIN

Query Match 100.0%; Score 26; DB 14; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;  
 Matches 26; Conservative 0; Mismatches 0;  
 QY 1 TCAACTTATGCCGCGTTGTACAGAC 26  
 |||||||  
 Db 63 TCAACTTATGCCGCGTTGTACAGAC 88

## RESULT 2

BE140710/c  
 LOCUS BE140710 229 bp mRNA linear EST 21-JUN-2000  
 DEFINITION CM0-HT0016-140699-008 HT0016 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BE140710  
 VERSION BE140710.1 GI:8603431  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 229)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 PUBLISHED

COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=CM0-HT0016-140699-008&tl3=1999-06-14&tl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 7  
 High quality sequence stop: 229.  
 Location/Qualifiers  
 1. .229  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0016"  
 /note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;

## FEATURES

source Location/Qualifiers  
 1. .229  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0016"  
 /note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;

## ORIGIN

Query Match 100.0%; Score 26; DB 10; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCAACTTATGCCGCGTTGTACAGAC 26  
 |||||||  
 Db 143 TCAACTTATGCCGCGTTGTACAGAC 118

## RESULT 3

BE159476  
 LOCUS BE159476 256 bp mRNA linear EST 21-JUN-2000  
 DEFINITION MRO-HT0407-100300-012-d07 HT0407 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BE159476  
 VERSION BE159476.1 GI:8622210  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 256)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed

## REFERENCE

## AUTHORS

1 (bases 1 to 256)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

## JOURNAL

## MEDLINE

## FUMED

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=MR0-HT0407-100300-012-d07&tl3=2000-03-10&tl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 256.  
 Location/Qualifiers  
 1. .256  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0407"  
 /note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 100.0%; Score 26; DB 10; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query Match 100.0%; Score 26; DB 10; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 136 TCAACTTATGCGCGTTGTACAGAC 161

RESULT 4
BE242285 265 bp mRNA linear EST 03-OCT-2001
LOCUS TCAAP1B1562 Pediatric acute myelogenous leukemia cell (FAB M1)
DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1562, mRNA
sequence.
ACCESSION BE242285
VERSION BE242285
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 265)
AUTHORS Wei Y., Tsang Y.T.M., Mei G., Xu J.M., Ali-Osman Jr., F.R.,
Muzzay D., Bouck J., Gibbs R.A. and Margolin C.F.
PEDIATRIC LEUKEMIA cDNA SEQUENCING PROJECT
UNPUBLISHED (2000)
CONTACT: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@ccc.org
Citation: Carninci, P. and Hayashizaki, Y. High efficiency
full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Seq primer: M13 primer.
FEATURES
Location/Qualifiers
1..265
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloc="TCAAP1562"
/sex="male"
/tissue_type="leukopheresis"
/dev_stage="myeloid cell"
/lab_host="DH10B"
/cloc_lib="Pediatric acute myelogenous leukemia cell (FAB
M1) Baylor-HGSC project-TCAA"
/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo (dT) primer [5'GGAGGACTCGAGCGCGCCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGAGCTCGATCCGCGCGCGCAATAATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T.,
Itoh M., Nagaoka S., Sasaki, Okazaki Y., Muramatsu M.,
Schneider C., Hayashizaki Y., High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 26, 1997)"
ORIGIN
Query Match 100.0%; Score 26; DB 10; Length 265;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 167 TCAACTTATGCGCGTTGTACAGAC 192

RESULT 5
AA360953 289 bp mRNA linear EST 21-APR-1997
LOCUS EST70157 T-cell lymphoma Homo sapiens cDNA 5' end similar to major
DEFINITION histocompatibility complex, class II antigen, alpha chain
(GI:X03100), mRNA sequence.
ACCESSION AA360953
VERSION AA360953.1 GI:20132273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 289)
AUTHORS Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A.,
Bult C.J., Lee N.H., Kirkness B.F., Weinstock K.G., Gocayne J.D.,
White O., Sutton G., Blake J.A., Brandon R.C., Man-Wai C.,
Clayton R.A., Cline T.R., Cotton M.D., Earle-Hughes J., Fine L.D.,
Fitzgerald L.M., Fitchugh W.M., Fritchman J.L., Geoghagen N.S.,
Glodek A., Gnehm C.L., Hanna M.C., Hedblom E., Hinkle P.S.Jr.,
Kelley J.M., Kelley J.C., Li L.-L., Marmaros S.M., Merrick J.M.,
Moreno-Palancas R.F., McDonald L.A., Nguyen D.R., Pelligrino S.M.,
Phillips C.A., Ryder S.E., Scott J.L., Saudek D.M., Shirley R.,
Small K.V., Spriggs T.A., Utterback T.R., Weidman J.F., Li Y.,
Bednarek D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J.,
Dimke D., Feng D.-F., Ferrie A., Fischer C., Hastings G.A.,
He W.W., Hu J.S., Greene J.M., Gruber J., Hudson P., Kim A.K.,
Kozak D.L., Kunsch C., Hungjun J., Li H., Meissner P.S., Olsen H.,
Raymond L., Wei Y.F., Wing J., Xu C., Yu G.B., Ruben S.M.,
Dillon P.J., Fannon M.R., Rosen C.A., Haseltine W.A., Fields C.,
Fraser C.M. and Venter J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL NATURE
MEDLINE 96036280
PUBMED 7566098
COMMENT Other ESTs: THCL72266
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/cdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..289
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):165076"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="T-cell lymphoma"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match 100.0%; Score 26; DB 9; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26
Db 142 TCAACTTATGCGCGTTGTACAGAC 167

RESULT 6
CD706205 290 bp mRNA linear EST 25-JUN-2003
LOCUS CD706205
DEFINITION EST22732 human nasopharynx Homo sapiens cDNA, tRNA sequence.

```

```

ACCESSION      CD706205
VERSION        CD706205.1  GI:32236835
SOURCE        EST.
ORGANISM      Homo sapiens (human)
KEYWORDS      Homo sapiens
REFERENCE      1. (bases 1 to 290)
AUTHORS      Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
              Zeng,Y.-X.
TITLE        Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL      Unpublished (2003)
COMMENT      Contact: Yixin Zeng
              Cancer Center
              Sun Yat-sen University
              651 Dongfeng Road East, Guangzhou 510060, China
              Tel: 86-1380-9770-743
              Fax: 86-20-8775-4506
              Email: yxzeng@gzsums.edu.cn.
              Location/Qualifiers
                1..290
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /issue_type="normal nasopharynx"
                  /clone_lib="human nasopharynx"
                  /note="ESTs generated from a normal nasopharynx cDNA
                    library from southern Chinese"

FEATURES             source
  source
    Query Match      100.0%; Score 26; DB 14; Length 290;
    Best Local Similarity 100.0%; Pred. No. 0.13;
    Matches          26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 TCACCTTATGCGCGTTGTACAGAC 26
180 TCACCTTATGCGCGTTGTACAGAC 205

CD694567
EST11090 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION      CD694567
VERSION        CD694567.1  GI:32219338
SOURCE        EST.
ORGANISM      Homo sapiens (human)
KEYWORDS      Homo sapiens
REFERENCE      1. (bases 1 to 295)
AUTHORS      Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
              Zeng,Y.-X.
TITLE        Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL      Unpublished (2003)
COMMENT      Contact: Yixin Zeng
              Cancer Center
              Sun Yat-sen University
              651 Dongfeng Road East, Guangzhou 510060, China
              Tel: 86-1380-9770-743
              Fax: 86-20-8775-4506
              Email: yxzeng@gzsums.edu.cn.
              Location/Qualifiers
                1..295
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /issue_type="normal nasopharynx"
                  /clone_lib="human nasopharynx"
                  /note="ESTs generated from a normal nasopharynx cDNA
                    library from southern Chinese"

FEATURES             source
  source
    Query Match      100.0%; Score 26; DB 14; Length 290;
    Best Local Similarity 100.0%; Pred. No. 0.13;
    Matches          26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 TCACCTTATGCGCGTTGTACAGAC 26
180 TCACCTTATGCGCGTTGTACAGAC 205

CD694567
EST11090 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION      CD694567
VERSION        CD694567.1  GI:32219338
SOURCE        EST.
ORGANISM      Homo sapiens (human)
KEYWORDS      Homo sapiens
REFERENCE      1. (bases 1 to 295)
AUTHORS      Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
              Zeng,Y.-X.
TITLE        Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL      Unpublished (2003)
COMMENT      Contact: Yixin Zeng
              Cancer Center
              Sun Yat-sen University
              651 Dongfeng Road East, Guangzhou 510060, China
              Tel: 86-1380-9770-743
              Fax: 86-20-8775-4506
              Email: yxzeng@gzsums.edu.cn.
              Location/Qualifiers
                1..295
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /issue_type="normal nasopharynx"
                  /clone_lib="human nasopharynx"
                  /note="ESTs generated from a normal nasopharynx cDNA
                    library from southern Chinese"

FEATURES             source
  source
    Query Match      100.0%; Score 26; DB 14; Length 290;
    Best Local Similarity 100.0%; Pred. No. 0.13;
    Matches          26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 TCACCTTATGCGCGTTGTACAGAC 26
180 TCACCTTATGCGCGTTGTACAGAC 205

```

```

Query Match      100.0%; Score 26; DB 14; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches          26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 TCACCTTATGCGCGTTGTACAGAC 26
171 TCACCTTATGCGCGTTGTACAGAC 196

RESULT 8
BM772902
LOCUS      K-EST0057161 S1SNUS Homo sapiens cDNA clone S1SNUS-8-C07 5', mRNA
DEFINITION sequence.
ACCESSION  BM772902
VERSION    BM772902.1  GI:19102517
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 335)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 8 Row: C Column: 07
            High quality sequence stop: 335.
            Location/Qualifiers
              1..335
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S1SNUS-8-C07"
                /sex="F"
                /issue_type="Ascites"
                /cell_type="lymphoblast-like"
                /cell_line="SNU-5"
                /lab_host="Top10F"
                /clone_lib="S1SNUS"
                /note="Organ: Stomach; Vector: pCNS; Site: 1: EcoRI;
                  Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
                  bacterial alkaline phosphatase (BAP) and then decapped
                  with tobacco acid pyrophosphatase (TAP). The decapped
                  intact mRNA was ligated with DNA-RNA linker including EcoR
                  I site by treatment of T4 RNA ligase and the first strand
                  cDNA was synthesized from oligo dt-selected mRNA by
                  priming with dt-tailed vector. The cDNA vector was
                  adjusted to have about 60nt. The cDNA after digestion of
                  circularized with E. coli DNA ligase after digestion of
                  EcoRI which site is also included in vector. An RNA strand
                  converted to a DNA strand by Okayama-Berg method. The
                  obtained cDNA vectors were used for transformation of
                  competent cells E. coli Top10F by electroporation method.
                  The cDNA libraries constructed by this method are
                  full-length enriched cDNA library."

FEATURES             source
  source
    Query Match      100.0%; Score 26; DB 12; Length 335;
    Best Local Similarity 100.0%; Pred. No. 0.13;
    Matches          26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 TCACCTTATGCGCGTTGTACAGAC 26
191 TCACCTTATGCGCGTTGTACAGAC 216

```

RESULT 9  
AA323639  
LOCUS  
DEFINITION  
EST26729 Cerebellum II Homo sapiens cDNA 5' end similar to major  
histocompatibility complex, class II antigen, alpha chain  
(GB:X03100), mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)

REFERENCE  
AUTHORS  
1 (bases 1 to 338)  
Adams M.D., Kerlavage A.R., Fleischmann R.D., Fulton R.A.,  
Bult C.J., Lee N.H., Kirkness E.F., Weinstock K.G., Gocayne J.D.,  
White O., Sutton G., Blake J.A., Brandon R.C., Man-Wai C.,  
Clayton R.A., Cline T.R., Cotton M.D., Earle-Hughes J., Fine L.D.,  
Fitzgerald L.M., Fitzhugh W.M., Fritchman J.L., Geoghegan N.S.,  
Glodek A., Gnehm C.B., Hanna M.C., Hedblom E., Hinkle P.S. Jr.,  
Kelley J.M., Kelley J.C., Liu L.-L., Marmaros S.M., Merrick J.M.,  
Moreno-Palancas R.P., McDonald L.A., Nguyen D.T., Pelligrino S.M.,  
Phillips C.A., Ryder S.E., Scott J.L., Saudek D.M., Shirley R.,  
Small K.V., Spriggs T.A., Utterback T.R., Weidman J.F., Li Y.,  
Bednarek D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J.,  
Dimke D., Feng D.-F., Fertie A., Fischer C., Hastings G.A.,  
He W.W., Hu J.S., Greene J.M., Gruber J., Hudson P., Kim A.K.,  
Kozak D.L., Kunsch C., Hungjun J., Li H., Meisner P.S., Olsen H.,  
Raymond L., Wei F.F., Wang J., Xu C., Yu G.L., Ruben S.M.,  
Dillon P.J., Fannon M.R., Rosen C.A., Haseltine W.A., Fields C.,  
Fraser C.M. and Venter J.C.

TITLE  
Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence

JOURNAL  
NATURE 377 (6547 Suppl), 3-174 (1995)

MDLINE  
95026280

PUBMED  
7566098

COMMENT  
Other ESTs: THC172266  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. .338  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):124435"  
/db\_xref="taxon:9606"  
/tissue\_type="cerebellum"  
/dev\_stage="adult"  
/clone\_lib="Cerebellum II"  
/notes="Organ: brain; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI"

ORIGIN  
Query Match 100.0%; Score 26; DB 9; Length 338;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTATCCCGCGTTGTACAGAC 26  
|||||  
DB 188 TCAACTATCCCGCGTTGTACAGAC 213  
|||||

RESULT 10

T94759  
LOCUS  
DEFINITION  
Yes7f12.r1 Stratagene lung (#937210) Homo sapiens cDNA clone  
IMAGE:119951 5' similar to GB:K01506 HLA CLASS II  
HISTOCOMPATIBILITY ANTIGEN, D2(1) ALPHA CHAIN (HUMAN);, mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)

REFERENCE  
AUTHORS  
1 (bases 1 to 366)  
Hillier L., Lennon G., Becker M., Bonaldo M.F., Chiapelli B.,  
Chissoe S., Dietrich N., DuBuque T., Favello A., Gish W.,  
Hawkins M., Hultman M., Kucaba T., Lacy M., Le M., Le N.,  
Mards E., Moore B., Morris M., Parsons J., Prange C., Rifkin L.,  
Rohlfing T., Schellenberg K., Soares M.B., Tan F., Thierry-Mieg J.,  
Trevasakis E., Underwood K., Wohldmann P., Waterston R., Wilson R.  
and Warra M.

TITLE  
Generation and analysis of 280,000 human expressed sequence tags

JOURNAL  
Genome Res. 6 (9), 807-828 (1996)

MDLINE  
97044478

PUBMED  
8889549

COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1152  
High quality sequence stops: 251 Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 1152 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 251.  
Location/Qualifiers  
1. .366  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:488240"  
/db\_xref="taxon:9606"  
/clone="IMAGE:119951"  
/sex="male"  
/dev\_stage="72 years"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/clone\_lib="Stratagene lung (#937210)"  
/notes="Organ: lung; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT, normal lung. Average insert size: 1.0 kb;  
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGCAGG  
3' -3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTT 3'"

ORIGIN  
Query Match 100.0%; Score 26; DB 14; Length 366;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTATCCCGCGTTGTACAGAC 26  
|||||  
DB 61 TCAACTATCCCGCGTTGTACAGAC 86  
|||||

RESULT 11

LOCUS  
DEFINITION  
CB267061  
1005967 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
cDNA 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
EST.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 05:13:19 ; Search time 604.235 Seconds  
(without alignments)  
1434.641 Million cell updates/sec

Title: US-09-877-819b-38

Perfect score: 20

Sequence: 1 ggctttggagcgcgtctttaa 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 180 summaries

Database :

GenBank:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.scs.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.to.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rcd.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	158412	2	AP001134 Homo sapi
2	18	90.0	2144	9	BC034803 Homo sapi
3	18	90.0	132953	9	AC024221 Homo sapi
4	18	90.0	161087	2	AC018850 Homo sapi
5	18	90.0	170067	2	AC124946 Homo sapi
6	18	90.0	188526	2	AC012342 Homo sapi
7	18	90.0	190943	2	AC130424 Homo sapi
8	18	90.0	328187	2	AC117393 Homo sapi
9	17.4	87.0	216198	2	AC109271 Mus muscu
10	16.8	84.0	30424	9	AC021658 Homo sapi
11	16.8	84.0	42775	1	ATVIRB
12	16.8	84.0	65525	9	AL161904
13	16.8	84.0	72440	2	AC113794_5
14	16.8	84.0	92800	9	AC004486_0
15	16.8	84.0	110000	2	BX546456_0
16	16.8	84.0	119790	9	HS01989D7
17	16.8	84.0	121866	2	HSJ322517
18	16.8	84.0	123066	2	AC140029
19	16.8	84.0	136384	2	AL1365221
20	16.8	84.0	153920	2	AC011723
21	16.8	84.0	160659	2	AC036235
22	16.8	84.0	163772	9	AC093165
23	16.8	84.0	163976	2	AC102004
24	16.8	84.0	165935	9	BX537254
25	16.8	84.0	166098	9	AL1356004
26	16.8	84.0	169788	2	AC113871
27	16.8	84.0	175588	9	AL139326
28	16.8	84.0	176182	2	AC137453
29	16.8	84.0	178505	2	AL611965
30	16.8	84.0	183379	2	AC121008
31	16.8	84.0	187564	9	CNS0000E
32	16.8	84.0	194140	1	AF242881
33	16.8	84.0	203047	10	AL928922
34	16.8	84.0	216397	2	AC119834
35	16.8	84.0	228273	2	AC111408
36	16.8	84.0	235971	2	AC124950
37	16.8	84.0	237657	2	AC093962
38	16.8	84.0	242569	2	AC122861
39	16.8	84.0	246789	2	AC119580
40	16.8	84.0	251516	2	AC130168
41	16.8	84.0	252646	2	AC137013
42	16.8	84.0	294307	2	AC095267
43	16.4	82.0	488	5	AF292650
44	16.4	82.0	512	11	BV004810
45	16.4	82.0	44557	1	AF305077
46	16.4	82.0	69161	1	AC101153
47	16.4	82.0	69161	2	AC101153
48	16.4	82.0	72887	2	AC017882
49	16.4	82.0	117482	9	AC105028
50	16.4	82.0	124230	10	AC005259
51	16.4	82.0	143180	2	AC120596
52	16.4	82.0	160931	3	AC093498
53	16.4	82.0	161841	3	AC093498
54	16.4	82.0	162063	9	AC135279
55	16.4	82.0	169869	2	AC005867
56	16.4	82.0	176121	9	AC130111
57	16.4	82.0	176932	9	AL157884
58	16.4	82.0	181121	10	AC058789
59	16.4	82.0	184263	9	AC004805
60	16.4	82.0	186451	10	AC122211
61	16.4	82.0	193743	2	AC125024
62	16.4	82.0	195988	2	AC141493
63	16.4	82.0	200207	2	AC007867
64	16.4	82.0	208030	10	AL672276
65	16.4	82.0	215532	9	AC010883

Pred. No. is the number of results predicted by chance to have a

66 16.4 82.0 219239 2 AC123661 Mus muscu  
 67 16.4 82.0 224697 2 AC147234 Mus muscu  
 68 16.4 82.0 234399 2 AC095649 Rattus no  
 69 16.4 82.0 236193 2 AC094506 Rattus no  
 70 16.4 82.0 303325 2 AC003538 Drosophil  
 71 16.4 82.0 307120 1 CNEPAX03  
 72 16.4 82.0 349980 6 AX041919 Sequence  
 73 16 80.0 258 6 AX041919 Sequence  
 74 16 80.0 258 6 AX041919 Sequence  
 75 16 80.0 1384 1 HPV418316  
 76 16 80.0 1545 1 AY155322  
 77 16 80.0 1545 1 AY155322  
 78 15.8 80.0 195290 9 AL354795  
 79 15.8 80.0 648 11 BV017308  
 80 15.8 79.0 648 6 AX790657 Sequence  
 81 15.8 79.0 673 6 AX789789 Sequence  
 82 15.8 79.0 759 6 AX654628 Sequence  
 83 15.8 79.0 1024 6 AX789421 Sequence  
 84 15.8 79.0 1024 6 AX789421 Sequence  
 85 15.8 79.0 1024 6 AX789421 Sequence  
 86 15.8 79.0 1071 5 AF549127 Cottus ba  
 87 15.8 79.0 1072 5 AF549120 Cottus co  
 88 15.8 79.0 1072 5 AF549119 Cottus co  
 89 15.8 79.0 1072 5 AF549145 Cottus co  
 90 15.8 79.0 1072 5 AF549112 Cottus ca  
 91 15.8 79.0 1073 5 AF549113 Cottus ca  
 92 15.8 79.0 1073 5 AF549115 Cottus rh  
 93 15.8 79.0 1079 5 AF549110 Cottus ca  
 94 15.8 79.0 1080 5 AF549110 Cottus ca  
 95 15.8 79.0 1080 5 AF549114 Cottus ca  
 96 15.8 79.0 1080 5 AF549114 Cottus rh  
 97 15.8 79.0 1080 5 AF549121 Cottus ca  
 98 15.8 79.0 1080 5 AF549122 Cottus ca  
 99 15.8 79.0 1080 5 AF549125 Cottus ba  
 100 15.8 79.0 1080 5 AF549126 Cottus ba  
 101 15.8 79.0 1080 5 AF549128 Cottus cf  
 102 15.8 79.0 1080 5 AF549129 Cottus cf  
 103 15.8 79.0 1080 5 AF549131 Cottus cf  
 104 15.8 79.0 1080 5 AF549132 Cottus cf  
 105 15.8 79.0 1080 5 AF549133 Cottus cf  
 106 15.8 79.0 1080 5 AF549134 Cottus cf  
 107 15.8 79.0 1080 5 AF549135 Cottus cf  
 108 15.8 79.0 1080 5 AF549136 Cottus cf  
 109 15.8 79.0 1080 5 AF549137 Cottus cf  
 110 15.8 79.0 1080 5 AF549138 Cottus cf  
 111 15.8 79.0 1080 5 AF549139 Cottus cf  
 112 15.8 79.0 1080 5 AF549140 Cottus cf  
 113 15.8 79.0 1080 5 AF549141 Cottus cf  
 114 15.8 79.0 1080 5 AF549144 Cottus cf  
 115 15.8 79.0 1080 5 AF549158 Cottus hy  
 116 15.8 79.0 1080 5 AF549159 Cottus hy  
 117 15.8 79.0 1080 5 AF549160 Cottus hy  
 118 15.8 79.0 1140 5 AF549161 Cottus hy  
 119 15.8 79.0 1141 5 AY050723 Caranx ru  
 120 15.8 79.0 1141 5 AY116342 Adysocott  
 121 15.8 79.0 1141 5 AY116343 Asprocott  
 122 15.8 79.0 1141 5 AY116344 Asprocott  
 123 15.8 79.0 1141 5 AY116345 Asprocott  
 124 15.8 79.0 1141 5 AY116347 Cottinell  
 125 15.8 79.0 1141 5 AY116349 Limnococt  
 126 15.8 79.0 1141 5 AY116350 Limnococt  
 127 15.8 79.0 1141 5 AY116351 Limnococt  
 128 15.8 79.0 1141 5 AY116352 Limnococt  
 129 15.8 79.0 1141 5 AY116355 Comephoru  
 130 15.8 79.0 1141 5 AY116356 Comephoru  
 131 15.8 79.0 1141 5 AY116357 Batrachoc  
 132 15.8 79.0 1141 5 AY116358 Batrachoc  
 133 15.8 79.0 1141 5 AY116359 Cottocome  
 134 15.8 79.0 1141 5 AY116360 Cottocome  
 135 15.8 79.0 1141 5 AY116361 Cottus xe  
 136 15.8 79.0 1141 5 AY116362 Paracottu  
 137 15.8 79.0 1141 5 AY116363 Cottus ba  
 138 15.8 79.0 1141 5 AY116367 Cottus sl

139 15.8 79.0 1170 6 BD092631 Identific  
 140 15.8 79.0 1471 3 AY052063 Drosophil  
 141 15.8 79.0 1784 8 AY156979 Zea mays  
 142 15.8 79.0 1851 6 AR396960 Sequence  
 143 15.8 79.0 4400 6 AR227535 Sequence  
 144 15.8 79.0 4761 8 SFADATITCP  
 145 15.8 79.0 8546 6 BD193552  
 146 15.8 79.0 8946 1 AF527793  
 147 15.8 79.0 13695 1 AR000586  
 148 15.8 79.0 14987 1 AC015466  
 149 15.8 79.0 16090 9 AL354812 Human DNA  
 150 15.8 79.0 21522 1 AC008734  
 151 15.8 79.0 26321 3 CEC08511  
 152 15.8 79.0 37386 3 CBRG43L07  
 153 15.8 79.0 40915 8 SPAC4G9  
 154 15.8 79.0 41645 8 SPAC4G9  
 155 15.8 79.0 42544 8 SPAC4G9  
 156 15.8 79.0 45974 8 SPAC4G9  
 157 15.8 79.0 55802 2 AC044912  
 158 15.8 79.0 59412 8 NCB7F21  
 159 15.8 79.0 60919 2 AC017150  
 160 15.8 79.0 61897 10 AL844223  
 161 15.8 79.0 68769 9 AC006027  
 162 15.8 79.0 72327 3 AC004713  
 163 15.8 79.0 73257 2 AC101211  
 164 15.8 79.0 78062 10 AL844572  
 165 15.8 79.0 87350 2 AC009752  
 166 15.8 79.0 92141 2 AC021565  
 167 15.8 79.0 95747 9 AL358175  
 168 15.8 79.0 95983 9 AC004874  
 169 15.8 79.0 103313 5 AL954844  
 170 15.8 79.0 103992 9 AL161736  
 171 15.8 79.0 104802 9 AL161736  
 172 15.8 79.0 109210 2 AC144760  
 173 15.8 79.0 110000 2 AC092789  
 174 15.8 79.0 110492 2 AP004139  
 175 15.8 79.0 112292 9 HSJ238D15  
 176 15.8 79.0 113415 2 AP004860  
 177 15.8 79.0 114964 10 AL691507  
 178 15.8 79.0 125020 9 AF429315  
 179 15.8 79.0 135330 10 AL772165  
 180 15.8 79.0 135330 10 AL772165

## ALIGNMENTS

RESULT 1  
 AP001134 158412 bp DNA linear HTG 30-MAY-2000  
 LOCUS Homo sapiens chromosome 18 clone RP11-795E17 map 18q12, WORKING  
 DEFINITION DRAFT SEQUENCE, 23 unordered pieces.  
 ACCESSION AP001134.2 GI:8118426  
 VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.  
 1 (bases 1 to 158412)  
 Fattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,  
 Homo sapiens 158,412 genomic DNA of 18q12  
 Published Only in DataBase (2000)  
 2 (bases 1 to 158412)  
 Fattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,  
 Direct Submission  
 Submitted (04-FEB-2000) Masahira Fattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,  
 Japan [E-mail:fattori@gsc.riken.go.jp,  
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,

Fax: 81-42-778-9924)  
 On May 31, 2000 this sequence version replaced gi:6997816.  
 ----- Genome Center  
 Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: <http://hgp.gsc.riken.go.jp/>  
 Contact: [hattori@gsc.riken.go.jp](mailto:hattori@gsc.riken.go.jp)  
 ----- Project Information  
 Center project name: HumDraft18  
 Center clone name: RP11-785E17  
 ----- Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 142804 bases at least Q40  
 Consensus quality: 151103 bases at least Q30  
 Consensus quality: 154448 bases at least Q20  
 Insert size: 156212; sum-of-contigs  
 Quality coverage: 4.59x in Q20 bases; sum-of-contigs  
 -----

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 23680 contig of 23680 bp in length
23781 44779 contig of 20999 bp in length
44880 58102 contig of 13223 bp in length
58203 70318 contig of 12116 bp in length
70419 78570 contig of 8152 bp in length
78671 88336 contig of 9666 bp in length
88437 96171 contig of 7735 bp in length
96272 104394 contig of 8123 bp in length
104495 109550 contig of 5056 bp in length
109651 117133 contig of 7483 bp in length
117234 122483 contig of 5250 bp in length
122584 127553 contig of 4970 bp in length
127654 132039 contig of 4386 bp in length
132140 136162 contig of 3023 bp in length
136263 139017 contig of 3755 bp in length
139118 141986 contig of 2869 bp in length
142087 144886 contig of 2698 bp in length
144987 147785 contig of 2677 bp in length
147851 152517 contig of 1956 bp in length
152618 154387 contig of 1770 bp in length
154488 157068 contig of 2581 bp in length
157169 158412 contig of 1244 bp in length
Sequence updated (26-May-2000).

```

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 23680: contig of 23680 bp in length
23681 23780: gap of 100 bp
23781 44779: contig of 20999 bp in length
44780 44879: gap of 100 bp
44880 58102: contig of 13223 bp in length
58103 58202: gap of 100 bp
58203 70318: contig of 12116 bp in length
70319 70418: gap of 100 bp
70419 78570: contig of 8152 bp in length
78571 78670: gap of 100 bp
78671 88336: contig of 9666 bp in length
88337 88436: gap of 100 bp
88437 96171: contig of 7735 bp in length
96172 96272: gap of 100 bp

```

```

* 96272 104394: contig of 8123 bp in length
* 104395 104494: gap of 100 bp
* 104495 109550: contig of 5056 bp in length
* 109551 117133: contig of 7483 bp in length
* 117134 117233: gap of 100 bp
* 117234 122483: contig of 5250 bp in length
* 122484 122583: gap of 100 bp
* 122584 127553: contig of 4970 bp in length
* 127554 127653: gap of 100 bp
* 127654 132039: contig of 4386 bp in length
* 132040 132139: gap of 100 bp
* 132140 136162: contig of 3023 bp in length
* 136163 136262: gap of 100 bp
* 136263 139017: contig of 3755 bp in length
* 139018 139117: gap of 100 bp
* 139118 141986: contig of 2869 bp in length
* 141987 142086: gap of 100 bp
* 142087 144886: contig of 2698 bp in length
* 144887 144986: gap of 100 bp
* 144987 147784: contig of 2698 bp in length
* 147785 147850: gap of 100 bp
* 147851 149740: contig of 1956 bp in length
* 149741 149840: gap of 100 bp
* 149841 152517: contig of 2677 bp in length
* 152518 152617: gap of 100 bp
* 152618 154387: contig of 1770 bp in length
* 154388 154487: gap of 100 bp
* 154488 157068: contig of 2581 bp in length
* 157069 157168: gap of 100 bp
* 157169 158412: contig of 1244 bp in length.
FEATURES
      source
      1..158412
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="18"
      /map="18q12"
      /clone="RP11-785E17"
      1..23680
      /note="assembly_fragment"
      23781..44779
      /note="assembly_fragment"
      44880..58102
      /note="assembly_fragment"
      58203..70318
      /note="assembly_fragment"
      70419..78570
      /note="assembly_fragment"
      78671..88336
      /note="assembly_fragment"
      88437..96171
      /note="assembly_fragment"
      96272..104394
      /note="assembly_fragment"
      104495..109550
      /note="assembly_fragment"
      109651..117133
      /note="assembly_fragment"
      117234..122483
      /note="assembly_fragment"
      122584..127553
      /note="assembly_fragment clone_end:r7 vector_side:left"
      127654..132039
      /note="assembly_fragment"
      132140..136162
      /note="assembly_fragment"
      136263..139017
      /note="assembly_fragment"
      139118..141986
      /note="assembly_fragment"
      142087..144886
      /note="assembly_fragment"

```

```

misc_feature 144987..147684
/note="assembly_fragment"
misc_feature 14785..149740
/note="assembly_fragment"
misc_feature 14941..152517
/note="assembly_fragment"
misc_feature 152618..154387
/note="assembly_fragment"
misc_feature 154488..157068
/note="assembly_fragment"
misc_feature 157169..158412
/note="assembly_fragment"

ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 158412;
Best Local Similarity 95.0%; Pred. No. 47;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTTAA 20
|||||
Db 74790 GGCTTTGGAGCGCTCTTTAA 74809

RESULT 2
BC034803
LOCUS BC034803 2144 bp mRNA linear PRI 26-JUL-2002
DEFINITION Homo sapiens, clone IVAGR:4821395, mRNA.
ACCESSION BC034803
VERSION BC034803.1 GI:21361560
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2144)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs@mail.nih.gov
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Format: 71 Row: c Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Excellent library of
origin.

FEATURES
Location/Qualifiers
1..2144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4821395"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

misc_feature 144987..147684
/note="assembly_fragment"
misc_feature 14785..149740
/note="assembly_fragment"
misc_feature 14941..152517
/note="assembly_fragment"
misc_feature 152618..154387
/note="assembly_fragment"
misc_feature 154488..157068
/note="assembly_fragment"
misc_feature 157169..158412
/note="assembly_fragment"

ORIGIN
Query Match 90.0%; Score 18; DB 9; Length 2144;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTTGGAGCGCTCTTT 18
|||||
Db 403 GGCTTTGGAGCGCTCTTT 420

RESULT 3
AC024221
LOCUS AC024221 132953 bp DNA linear PRI 29-AUG-2002
DEFINITION Homo sapiens 3 BAC RP11-22705 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
ACCESSION AC024221
VERSION AC024221.23 GI:21206011
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 132953)
AUTHORS Muzny,D.M., Adams,C., Agio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaraturge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dint,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotte,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,K., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,B., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisseg,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Neelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Ogih,K., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,R., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergran,E.,
Sonaite,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.P., Zhou,J., Zorrilla,S., Naylor,S.I., Weinstein,G. and
Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 132953)
Worley,K.C.
Direct Submission
Submitted (28-FEB-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

```

```

REFERENCE
AUTHORS      3 (bases 1 to 132953)
TITLE        Worley,K.C.
JOURNAL      Direct Submission:
              Submitted (21-MAY-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
REFERENCE
AUTHORS      4 (bases 1 to 132953)
TITLE        Worley,K.C.
JOURNAL      Direct Submission
              Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
REFERENCE
AUTHORS      5 (bases 1 to 132953)
TITLE        Worley,K.C.
JOURNAL      Direct Submission
              Submitted (29-AUG-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
COMMENT      On May 25, 2002 this sequence version replaced gi:20986394.
              INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
              gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
  STSs are identified using ePCR (Genome Res. 7:541-550) searches
  of a local database that includes entries from dbSTS, GDB, and
  local mapping efforts.
  Repeats are identified using RepeatMasker (A. Smit and P. Green,
  unpublished.) for Human and Mouse sequences.
  Genes and Region of sequence similarity are identified by BLAST
  (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
  EST and cDNA sequences. Genes demonstrate at least two exons
  flanked by consensus splice sites that maintained sequence
  continuity across the splice junctions. Sequences that are not
  identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES
source      Location/Qualifiers
            1..132953
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="3"
              /clone="RP11-227J5"
            1..2036
              /note="overlaps bases 122666..124701 of clone AC079594"
              /function="clone overlap"
              /rpt_family="AluJo"
              995..1028
                /rpt_family="AT-rich"
                complement(1298..1602)
                  /rpt_family="AluJo"

misc_feature
            1..2036
              /note="overlaps bases 122666..124701 of clone AC079594"
              /function="clone overlap"
              /rpt_family="AluJo"
              995..1028
                /rpt_family="AT-rich"
                complement(1298..1602)
                  /rpt_family="AluJo"

repeat_region
            90.08; Score 18; DB 9; Length 132953;
            Best local Similarity 100.0%; Pred. NO. 79;
            Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

repeat_region
            complement(1889..2090)
              /rpt_family="AluJo"
            2770..2808
              /rpt_family="(TTTA)n"
            complement(2809..3094)
              /rpt_family="AluJb"
            3313..3372
              /rpt_family="MER2"
            complement(3379..3435)
              /rpt_family="AluJ/FLAX"
            3576..3749
              /rpt_family="L2"
            3648..3796
              /standard_name="15013"
            3648..3796
              /standard_name="69663"
            4160..4247
              /rpt_family="L1p"
            4246..49783
              /rpt_family="L1PA6"
            8111..8390
              /standard_name="152451"
            complement(10336..10612)
              /rpt_family="MIR"
            11396..11451
              /rpt_family="MES5B"
            complement(11458..12087)
              /rpt_family="L1MA2"
            12834..12866
              /rpt_family="(TGAA)n"
            13142..13161
              /rpt_family="(A)n"
            14076..14103
              /rpt_family="AT-rich"
            14383..14427
              /rpt_family="AT-rich"
            15029..15069
              /rpt_family="MIR"
            complement(15234..16523)
              /rpt_family="AluJb"
            17157..17435
              /standard_name="134217"
            17723..17856
              /rpt_family="L1M4"
            complement(17858..17989)
              /rpt_family="AluJb"
            complement(18232..18755)
              /rpt_family="MER1A"
            18951..19215
              /rpt_family="AluJb"
            complement(21057..21293)
              /rpt_family="L2"
            21326..21852
              /rpt_family="MLTIF1"
            22483..22511
              /rpt_family="AT-rich"
            23197..24501
              /rpt_family="L1MC4"
            complement(24573..24624)
              /rpt_family="MIR"
            24792..24810
              /rpt_family="MLT1K"
            complement(24811..25083)
              /rpt_family="AluSx"
            25084..25136
              /rpt_family="MLT1K"
            25735..26018
              /rpt_family="MLT1K"

```

Qy 1 GGCTTGAGCGCTCTT 18

```

Db      95222  GCGTTGGAGCGCTCTT  95239
|||||
RESULT 4
AC018850/c
LOCUS   AC018850
DEFINITION Homo sapiens chromosome 3 clone RP11-416D22 map 3, WORKING DRAFT
SEQUENCE, 21 unordered pieces.
ACCESSION AC018850
VERSION   AC018850.3 GI:8096873
KEYWORDS  HTG; HTGS_PHASE; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161087)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 3, clone RP11-416D22
Unpublished
2 (bases 1 to 161087)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,P.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGuirk,A., McKernan,K.,
Meldrim,J., Menius,L., Morrow,J., Naylor,G., Norman,C.H.,
O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N.,
Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,C., Tirrell,A.,
Vassiliev,E., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
and Zody,M.
Direct Submission
Submitted (21-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced gi:6716059.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WTRR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1824
Center clone name: 416 D.22
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142691 bases at least Q40
Consensus quality: 150851 bases at least Q30
Consensus quality: 154724 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 159087; sum-of-contigs
Quality coverage: 3.3 in Q20 bases; agarose-fp
Quality coverage: 3.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1096: contig of 1096 bp in length  
1196: gap of 100 bp  
2718: contig of 1522 bp in length  
2719: gap of 100 bp  
2818: gap of 100 bp  
3925: contig of 1107 bp in length  
3926: gap of 100 bp  
4025: gap of 100 bp  
5054: contig of 1029 bp in length  
5154: gap of 100 bp  
7181: contig of 2027 bp in length  
7281: gap of 100 bp  
11155: contig of 3874 bp in length  
11255: gap of 100 bp  
14790: contig of 3535 bp in length  
14791: gap of 100 bp  
14891: gap of 100 bp  
18942: contig of 4052 bp in length  
19043: gap of 100 bp  
23097: contig of 4055 bp in length  
23198: gap of 100 bp  
31007: contig of 7810 bp in length  
31008: gap of 100 bp  
31176: contig of 6069 bp in length  
37276: gap of 100 bp  
44520: contig of 7244 bp in length  
44620: gap of 100 bp  
52862: contig of 8242 bp in length  
52962: gap of 100 bp  
60848: contig of 7886 bp in length  
60948: gap of 100 bp  
68864: contig of 7916 bp in length  
68964: gap of 100 bp  
75113: contig of 6749 bp in length  
75813: gap of 100 bp  
84405: contig of 8592 bp in length  
84505: gap of 100 bp  
96213: contig of 11708 bp in length  
96313: gap of 100 bp  
107681: contig of 11368 bp in length  
107781: gap of 100 bp  
133144: contig of 25363 bp in length  
133244: gap of 100 bp  
161087: contig of 27843 bp in length.

Location/Qualifiers  
1. .161087  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3"  
/clone="RP11-416D22"  
/clone\_lib="RPC1-11 Human Yale BAC"  
1. .1096  
/note="assembly\_fragment"  
1197. .2718  
/note="assembly\_fragment"  
2819. .3925  
/note="assembly\_fragment"  
4026. .5054  
/note="assembly\_fragment"  
5155. .7181  
/note="assembly\_fragment"  
7282. .11155  
/note="assembly\_fragment"  
11256. .14790  
/note="assembly\_fragment"  
14891. .18942  
/note="assembly\_fragment"  
19043. .23097  
/note="assembly\_fragment"  
23198. .31007  
/note="assembly\_fragment"  
31108. .37176  
/note="assembly\_fragment"  
37277. .44520  
/note="assembly\_fragment"

FEATURES  
source  
1. .161087  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3"  
/clone="RP11-416D22"  
/clone\_lib="RPC1-11 Human Yale BAC"  
1. .1096  
/note="assembly\_fragment"  
1197. .2718  
/note="assembly\_fragment"  
2819. .3925  
/note="assembly\_fragment"  
4026. .5054  
/note="assembly\_fragment"  
5155. .7181  
/note="assembly\_fragment"  
7282. .11155  
/note="assembly\_fragment"  
11256. .14790  
/note="assembly\_fragment"  
14891. .18942  
/note="assembly\_fragment"  
19043. .23097  
/note="assembly\_fragment"  
23198. .31007  
/note="assembly\_fragment"  
31108. .37176  
/note="assembly\_fragment"  
37277. .44520  
/note="assembly\_fragment"

misc\_feature  
1. .1096  
/note="assembly\_fragment"  
1197. .2718  
/note="assembly\_fragment"  
2819. .3925  
/note="assembly\_fragment"  
4026. .5054  
/note="assembly\_fragment"  
5155. .7181  
/note="assembly\_fragment"  
7282. .11155  
/note="assembly\_fragment"  
11256. .14790  
/note="assembly\_fragment"  
14891. .18942  
/note="assembly\_fragment"  
19043. .23097  
/note="assembly\_fragment"  
23198. .31007  
/note="assembly\_fragment"  
31108. .37176  
/note="assembly\_fragment"  
37277. .44520  
/note="assembly\_fragment"

```

/misc_feature /note="assembly_fragment"
44521..52862
/misc_feature /note="assembly_fragment"
52963..60848
/misc_feature /note="assembly_fragment"
60949..68864
/misc_feature /note="assembly_fragment"
68965..75713
/misc_feature /note="assembly_fragment"
75814..84405
/misc_feature /note="assembly_fragment"
84506..96213
/misc_feature /note="assembly_fragment"
96314..107681
/misc_feature /note="assembly_fragment
clone_end:SP6
vector side:right"
107782..133144
/misc_feature /note="assembly_fragment"
133245..161087
/misc_feature /note="assembly_fragment"

```

## ORIGIN

```

Query Match 90.0%; Score 18; DB 2; Length 161087;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCCTTTGGAGCGCTCTTT 18

Db 72742 GCCTTTGGAGCGCTCTTT 72725

## RESULT 5

AC124946/c

LOCUS

DEFINITION

AC124946

VERSION

HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

170067 bp DNA linear HTG 29-JUN-2002  
Homo sapiens clone RP11-416D22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 3  
unordered pieces.

AC124946 3 GI:21539690  
HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN.

Homo sapiens  
Homo sapiens (human)

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 170067)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,P.R., Allen,C.,  
Alsbrooks,S.L., Amaral,J., Bimonte,N., Binkley,M., Banks,T.,  
Barbata,J., Benton,J., Bimonte,N., Binkley,M., Banks,T.,  
Bouck,J., Bowie,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Pavy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flegg,N., Ford,J.J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,P., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvat,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Louseghed,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,K.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telifrod,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 170067)  
Worley,K.C.

Direct Submission  
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 170067)  
Worley,K.C.

Direct Submission  
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jun 23, 2002 this sequence version replaced gi:2153969046.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HEED

Center clone name: RP11-416D22

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 168450 bases at least Q40

Consensus quality: 168504 bases at least Q30

Consensus quality: 168514 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 3 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 15686: contig of 15686 bp in length

\* 15687 15786: gap of unknown length

\* 15787 61007: contig of 45221 bp in length

\* 61008 61107: gap of unknown length

\* 61108 170067: contig of 108960 bp in length.

## FEATURES

source

1..170067

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="RP11-416D22"

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 170067;

Best Local Similarity 100.0%; Pred. No. 80;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTTGGAGCGCTCTTT 18

Db 164737 GCCTTTGGAGCGCTCTTT 164720

```

RESULT 6
AC012342
LOCUS
DEFINITION
Homo sapiens chromosome 3 clone RP11-436F13 map 3, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
AC012342
AC012342.3 GI:8096867
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 188526)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Co-angelo,M., Collins,S., Collymore,A., B.,
Cooke,P., Dearellano,K., Dewar,X., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gager,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lenoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McSwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,K., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced GI:6454045.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1812
Center clone name: 436 F 13
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152028 bases at least Q40
Consensus quality: 172958 bases at least Q30
Consensus quality: 181270 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 186426; sum-of-contigs
Quality coverage: 4.8 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1030: contig of 1030 bp in length
* 1031 1130: gap of 100 bp
* 1131 2538: contig of 1408 bp in length
* 2539 2638: gap of 100 bp
* 2639 3766: contig of 1128 bp in length

```



misc\_feature 40262.48428  
 /note="assembly\_fragment"  
 misc\_feature 48529.61353  
 /note="assembly\_fragment"  
 misc\_feature 61454.72822  
 /note="assembly\_fragment"  
 misc\_feature 72923.85724  
 /note="assembly\_fragment"  
 misc\_feature 85825.96678  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:right  
 96779.112114  
 /note="assembly\_fragment"  
 misc\_feature 112215.129962  
 /note="assembly\_fragment"  
 clone\_end:17  
 vector\_side:left  
 130063.145946  
 /note="assembly\_fragment"  
 misc\_feature 146047.188526  
 /note="assembly\_fragment"

## ORIGIN

Query Match 90.0% Score 18; DB 2; Length 188526;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTGGAGCGCTCTTT 18  
 Db 44172 GGCTTGGAGCGCTCTTT 44189

## RESULT 7

AC130424  
 LOCUS  
 DEFINITION Homo sapiens chromosome 3 clone RP11-436F13, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 18 unordered pieces.  
 AC130424  
 VERSION AC130424.1 GI:22203181  
 KEYWORDS HTG; HTGS PHASE1.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 190943)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buha,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Dent,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,  
 Housi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,K., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseg,H.,  
 Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Maxwell,E., McLeod,W.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogrin,M., Okwuonu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,  
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,P., Telford,B., Thomas,N., Thomas,S.,  
 Umani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 190943)  
 Worley,K.C.  
 Direct Submission  
 Submitted (10-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 190943)  
 Worley,K.C.  
 Direct Submission  
 Submitted (08-JAN-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center Project name: HDBE  
 Center clone name: RP11-436F13  
 ----- Summary Statistics  
 Chemistry: Dye-terminator Big Dye: Infinity% of reads  
 Chemistry: Dye-terminator Big Dye: Infinity% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 176145 bases at least Q40  
 Consensus quality: 182574 bases at least Q30  
 Consensus quality: 185981 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 18 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 2258: contig of 2258 bp in length  
 2358: gap of unknown length  
 5105: contig of 2747 bp in length  
 5205: gap of unknown length  
 5206: contig of 2538 bp in length  
 7743: gap of unknown length  
 7744: contig of 2655 bp in length  
 10498: contig of 2097 bp in length  
 10598: gap of unknown length  
 12695: contig of 2097 bp in length  
 12796: gap of unknown length  
 14960: contig of 2165 bp in length  
 14961: gap of unknown length  
 15061: contig of 4507 bp in length  
 15067: gap of unknown length  
 19568: contig of 5456 bp in length  
 19668: gap of unknown length  
 25124: contig of 8056 bp in length  
 33279: gap of unknown length  
 33280: contig of 9270 bp in length  
 42649: gap of unknown length  
 42650: contig of 9270 bp in length

\* 42750 54103: contig of 11354 bp in length  
 \* 54104 54203: gap of unknown length  
 \* 54204 54461: contig of 11258 bp in length  
 \* 65461 65561: gap of unknown length  
 \* 65562 78517: contig of 12956 bp in length  
 \* 78518 78617: gap of unknown length  
 \* 78618 94021: contig of 15404 bp in length  
 \* 94022 94121: gap of unknown length  
 \* 94122 112067: contig of 17946 bp in length  
 \* 112068 112167: gap of unknown length  
 \* 112168 128112: contig of 15945 bp in length  
 \* 128113 128212: gap of unknown length  
 \* 128213 148347: contig of 20135 bp in length  
 \* 148348 148447: gap of unknown length  
 \* 148448 190943: contig of 42496 bp in length.

## FEATURES

source

Location/Qualifiers  
 ..190943  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP211-436F13"

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 190943;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTTTGAGCGCTCTTT 18

Db 37841 GCGTTTGAGCGCTCTTT 37858

## RESULT 8

AC117393

## LOCUS

DEFINITION Homo sapiens chromosome 3 clone RP11-372J9, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\* 17 unordered pieces.

AC117393 AC012057

AC117393.2 GI:20279204

HTG; HTGS\_PHASE1; HTGS\_ACTIVEPIN.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 328187)

## REFERENCE

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,P.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Garroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinch,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,I., Farraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gable,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,  
 Honsi,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L., Lewis,L.,  
 Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loussege,H.,  
 Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawliny,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,  
 Ocranbye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,T.,  
 Sodergren,E., Sonaitte,T., Sparks,A., Stanley,H., Stone,K., Tang,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

## Direct Submission

Unpublished  
 2 (bases 1 to 328187)

Worley,K.C.

## Direct Submission

Submitted (10-APR-2002)

Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 328187)

Worley,K.C.

## Direct Submission

Submitted (29-APR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On or before Apr 28, 2002 this sequence version replaced  
 GI:8072477, GI:20127689.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HDGP

Center clone name: RP11-372J9

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 345031 bases at least Q40

Consensus quality: 361075 bases at least Q30

Consensus quality: 371672 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 17 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 2010: contig of 2010 bp in length  
 2110: gap of unknown length  
 2111: contig of 2440 bp in length  
 4550: gap of unknown length  
 4551: contig of 2660 bp in length  
 7310: gap of unknown length  
 7311: contig of 2650 bp in length  
 7411: gap of unknown length  
 10061: contig of 3851 bp in length  
 10161: contig of 4624 bp in length  
 14011: gap of unknown length  
 14012: contig of 5138 bp in length  
 18735: gap of unknown length  
 18836: contig of 5996 bp in length  
 23974: gap of unknown length  
 24073: contig of 5996 bp in length  
 30070: gap of unknown length  
 30170: contig of 7470 bp in length  
 37739: gap of unknown length

\* 37740 46044: contig of 8305 bp in length  
 \* 46144: gap of unknown length  
 \* 46145 57813: contig of 11669 bp in length  
 \* 57814 75913: gap of unknown length  
 \* 75914 73711: contig of 15798 bp in length  
 \* 73712 73811: gap of unknown length  
 \* 73812 89444: contig of 15633 bp in length  
 \* 89445 110717: contig of 21173 bp in length  
 \* 110718 110817: gap of unknown length  
 \* 110818 131687: contig of 20870 bp in length  
 \* 131688 131787: gap of unknown length  
 \* 131788 165404: contig of 33617 bp in length  
 \* 165405 165505: gap of unknown length  
 \* 165506 328187: contig of 162683 bp in length.

## FEATURES

Location/Qualifiers  
 1..328187  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP11-3729"

## ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 328187;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGCTTTGGAGCGCTCTT 18  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 62499 GGCTTTGGAGCGCTCTT 62516

## RESULT 9

AC109271  
 LOCUS AC109271 216198 bp DNA linear HTG 11-DEC-2003  
 DEFINITION Mus musculus chromosome 1 clone RP23-353119 map 1, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 10 ordered pieces.  
 AC109271  
 VERSION AC109271.6 GI:39725831  
 KEYWORDS HTG; HTGS PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 216198)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 1, clone RP23-353119  
 Unpublished

## REFERENCE

## JOURNAL

## AUTHORS

2 (bases 1 to 216198)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouckgaiter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Glende,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Katat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
 Lander,T., Lehotzky,J., Levine,K., Liu,G., MacLean,C.,  
 Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,  
 McEwan,P., McKerran,K., Meldrum,J., Meneus,L., Mihova,T.,  
 Norman,C.H., O'Connor,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
 Retta,R., Rieback,W., Riley,R., Rise,C., Rogov,P., Roman,J.,  
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,  
 Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travis,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 216198)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavskiy,L., Bouckgaiter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., DeArillano,K., Faro,S.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Gardyna,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrum,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thoman,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travlers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (11-DEC-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 11, 2003 this sequence version replaced gi:38708121.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L20943  
 Center clone name: 353\_1\_19

## NOTE: This is a 'working draft' sequence. It currently

\* consists of 10 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 50562: contig of 50562 bp in length

\* 50563 50662: gap of 100 bp

\* 50663 91756: contig of 41094 bp in length

\* 91757 91856: gap of 100 bp

\* 91857 96746: contig of 4890 bp in length

\* 96747 96846: gap of 100 bp

\* 96847 135026: contig of 38180 bp in length

\* 135027 135126: gap of 100 bp

\* 135127 141481: contig of 6355 bp in length

\* 141482 141581: gap of 100 bp

\* 141582 145728: contig of 4147 bp in length

\* 145729 145828: gap of 100 bp

\* 145829 157908: contig of 12080 bp in length

\* 157909 158008: gap of 100 bp

\* 158009 185740: contig of 27732 bp in length

\* 185741 185840: gap of 100 bp

\* 185841 193106: contig of 7266 bp in length

\* 193107 193206: gap of 100 bp

\* 193207 216198: contig of 22992 bp in length.

Location/Qualifiers

1..216198  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="1"

## FEATURES

## SOURCE

/map="1"  
/clone="RP23-353119"  
/clone\_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 216198;  
Best Local Similarity 94.7%; Pred. No. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GCTTGGAGCCTCTTTAA 20  
|||||

Db 89456 GCTTGGAGCCTCTTTAA 89474  
|||||

RESULT 10  
AC021658/c  
LOCUS  
DEFINITION Homo sapiens 12 c174f5 complete sequence.  
AC021658 linear PRI 01-DEC-2000  
VERSION AC021658.18 GI:11496326  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (Bases 1 to 30424)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimave,K., Blankenburg,K., Bonnin,B., Bouck,J.,  
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleveland,C.D.,  
Cox,C., Coyne,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,  
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,  
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,  
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,  
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,F.,  
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homai,F.,  
Howard,S., Huber,J., Huijck,S., Hume,J., Ioshikhes,I., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,  
Lewis,L., Li,J., Li,Z., Licharge,O., Liek,C., Liu,J., Liu,W.,  
Loulseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,N., Mapua,P., Marondel,I., Martin,R.,  
Martindale,A., Martinez,E., Massey,B., Mawhinney,E., McLeod,M.P.,  
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,  
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,  
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
Cruh,M., Okwuon,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,  
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C.,  
Shoohatari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A.,  
Stanley,H., Stone,H., Tasey,J., Taylor,C., Taylor,T., Teifrod,B.,  
Tamerisa,K., Tang,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,  
Thomas,R., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,  
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,  
Washington,C., Watlington,S., Williams,G., Williamson,A.,  
Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Zhou,J.,  
Zorrilla,S., Zucherlapati,R. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 30424)  
Worley,K.C.  
Direct Submission  
Submitted (-9-JAN-2000) Human Genome Sequencing Center, Department

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
Worley,K.C.  
Direct Submission  
Submitted (01-DEC-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 1, 2000 this sequence version replaced gi11128254.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:

of STGs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse Sequences.  
Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----  
Contig length: 30424  
Phrap values in estimate: 30207  
Average error rate (BCM-Phrap estimate): 1.24133e-06  
Fraction of Phrap values less than 40 : 0.00218492  
Number of consensus changing edits: 0  
Number of N's in consensus : 0

----- Consensus changing edits ----- Edited+Context  
Position Original+Context

----- Distribution of Quality < 40 Bases -----

100
90
80
70
60
50
40
30
20

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

```

10|
0|
-----
5 10 15 20 25 30 35 40
Phrap Value Range

```

```

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCTTTGGAGCGCTCTTTAA 20
| | | | | | | | | | | | | | | | | |
Db 2468 GCCTTGGAGTGCCTCTTTAA 2449

```

## RESULT 11

```

ATVIRB 42775 bp DNA linear BCT 06-JUL-2002
Agrobacterium tumefaciens plasmid pTil5955.
LOCUS X06826 AP241250
DEFINITION X06826.2 GI:11124672
ACCESSION X06826.2
VERSION 1
KEYWORDS chaperone-like protein; CspA-like protein; endonuclease; IHP beta
subunit-like protein; insertion sequence IS66; insertion sequence
IS66-P; insertion sequence IS71; insertion sequence IS89-1; ORF1;
ORF2; ORF3; ORF4; ORF5; overlapping genes; pilin protein; single
strand DNA binding protein; TraA-like protein; TraB-like protein;
TraP-like protein; virA gene; virB operon; virB1 gene; virB10 gene;
virB11 gene; virB2 gene; virB3 gene; virB4 gene; virB5 gene; virB6
gene; virB7 gene; virB8 gene; virB9 gene; virC1 gene; virC2 gene;
virD1 gene; virD2 gene; virD3 gene; virD4 gene; virD5 gene; virE1
gene; virE2 gene; virE3 gene; virF gene; virG gene; virH2 gene;
virJ gene; virK gene.
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM Agrobacterium tumefaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1
AUTHORS Thompson,D.V., Melchers,L.S., Idler,K.B., Schilperoort,R.A. and
Hooijkaas,P.J.
TITLE Analysis of the complete nucleotide sequence of the Agrobacterium
tumefaciens virB operon
JOURNAL Nucleic Acids Res. 16 (10), 4621-4636 (1988)
MEDLINE 88247765
PUBMED 2837739
REFERENCE 2
AUTHORS Schrammeijer,B., Beijersbergen,A., Idler,K.B., Melchers,L.S.,
Thompson,D.V. and Hooijkaas,P.J.
TITLE Sequence analysis of the vir-region from Agrobacterium tumefaciens
octopine Ti plasmid pTil5955
J. Exp. Bot. 347, 1167-1169 (2000)
REFERENCE 3
AUTHORS Melchers,L.S.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1988) L.S. Melchers, Department of Plant
Molecular Biology, Biochemistry Lab., Leiden University,
Massenaarseweg 64, 2333 AL Leiden, The Netherlands
revised by [3]
REFERENCE 4
AUTHORS Thompson,D.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1990)
REMARK revised by [4]
REFERENCE 5
AUTHORS Schrammeijer,B.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2000) B. Schrammeijer, Molecular Genetics,
Institute of Molecular Plant Sciences, Wassenaarseweg 64, Leiden
2333 AL, NETHERLANDS
COMMENT On Nov 8, 2000 this sequence version replaced gi:39195.
Data kindly reviewed (21-OCT-1988) by Melchers L.S.
FEATURES
source
1..42775
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"
/plasmid="pTil5955"
1..201
/gene="virH2"
<1..201
/gene="virH2"
/codon_start=1
gene
CDS
Query Match 84.0%; Score 16.8; DB 9; Length 30424;
Best Local Similarity 90.0%; Pred. No. 3.8e-02;

```





```

repeat_region /note="MIR repeat: matches 9. .262 of consensus"
20507. .20998
repeat_region /note="98 copies 4 mer tott 62% conserved"
20644. .20985
repeat_region /note="171 copies 2 mer tt 63% conserved"
21047. .21242
repeat_region /note="AluSg/x repeat: matches 116. .305 of consensus"
22051. .22324
repeat_region /note="AluSg repeat: matches 21. .294 of consensus"
22594. .22789
repeat_region /note="L1PA13 repeat: matches 464. .4662 of consensus"
22790. .23099
repeat_region /note="AluSg repeat: matches 1. .309 of consensus"
23100. .24191
repeat_region /note="L1PA13 repeat: matches 4662. .5667 of consensus"
24192. .24496
repeat_region /note="AluSg repeat: matches 1. .301 of consensus"
24497. .24978
repeat_region /note="L1PA13 repeat: matches 5667. .6156 of consensus"
25899. .26269
repeat_region /note="THE1C repeat: matches 1. .371 of consensus"
26279. .26390
repeat_region /note="MIR repeat: matches 15. .121 of consensus"
26424. .26663
repeat_region /note="L1ME2 repeat: matches 5914. .6161 of consensus"
26673. .26705
repeat_region /note="11 copies 3 mer ttg 97% conserved"
26707. .26897
repeat_region /note="AluSg/x repeat: matches 88. .280 of consensus"
26985. .27172
repeat_region /note="L2 repeat: matches 2519. .2696 of consensus"
28665. .28835
repeat_region /note="MIR repeat: matches 1. .184 of consensus"
30034. .30169
repeat_region /note="MIR repeat: matches 28. .161 of consensus"
31438. .31887
repeat_region /note="match: GSS: Em-B16727"
3252. .35407
repeat_region /note="L1PA13 repeat: matches 3266. .6145 of consensus"
35408. .35718
repeat_region /note="AluY repeat: matches 1. .311 of consensus"
35719. .35755
repeat_region /note="L1PA13 repeat: matches 3230. .3266 of consensus"
35757. .35808
repeat_region /note="13 copies 4 mer tata 96% conserved"
35822. .36151
repeat_region /note="L1PA13 repeat: matches 2889. .3226 of consensus"
36159. .37610
repeat_region /note="L1PA12 repeat: matches 267. .1735 of consensus"
37588. .38264
repeat_region /note="L1PA12 repeat: matches 163. .575 of consensus"
38278. .38504
repeat_region /note="L1PA12 repeat: matches 1421. .1192 of consensus"
38716. .39771
repeat_region /note="L1M3 repeat: matches 4702. .5793 of consensus"
39771. .40068

```

Query Match 84.0%; Score 16.8; DB 9; Length 65825;  
 Best Local Similarity 90.0%; Pred. No. 3.8e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGTTGGAGCGCTCTTTAA 20  
 Db 55940 GCGTTGGAGCGCTCTTTGA 55959

RESULT 13  
 AC113794\_5  
 WPCOMMENT  
 Sequence split into 6 fragments LOCUS AC113794 Accession AC113794  
 Fragment Name Begin End  
 AC113794\_0 1 110000  
 AC113794\_1 100001 210000

AC113794\_2 200001 310000  
 AC113794\_3 300001 410000  
 AC113794\_4 400001 510000  
 AC113794\_5 500001 572440

Continuation (6 of 6) of AC113794 from base 500001 (AC113794 Rattus norvegicus clone CH2)

Query Match 84.0%; Score 16.8; DB 2; Length 72440;  
 Best Local Similarity 90.0%; Pred. No. 3.8e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGTTGGAGCGCTCTTTAA 20  
 Db 14631 GCGTTGGAGCGCTCTTTAA 14650

RESULT 14  
 AC004486/c

LOCUS  
 DEFINITION Homo sapiens 12q13 PAC RPC13-432118 (Roswell Park Cancer Institute Human PAC library) complete sequence.

ACCESSION AC004486  
 VERSION AC004486.1 GI:3776601

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Muzny, D., Arenson, A.D., Adams, C., Bunac, C., Carvelli, K., Chang, J., Chacko, J., Chen, J., Ding, Y., Dugan, S., Durbin, J., Forcum, J., Ganesh, R., Garcia, C., Goodman, M., Gorrell, J. H., Haywood, M., Hernandez, J., Jackson, L., Jin, S., Kampal, R., Karpathy, S., Kovar, C., Lau, S., Leal, B., Lee, E., Li, Y., Lichtarge, O., Liu, W., Logan, O., Lu, J., Ly, T., Marondel, I., Martinez, C., Merscher, S., Montgomery, K., Oswal, C., Perez, L., Rashid, N.D., Renault, B., Rowland, K., Savage, L., Scherer, S.E., Shen, H., Simon, M., Stovall, K., Timms, K.M., Todd, J., Vo, Q., Williamson, A., Worley, K.C., Yu, W., Kucherlapati, R., Nelson, D. and Gibbs, R.A.

Direct Submission

TITLE

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 92800)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (28-MAR-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 92800)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-1998) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 92800)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (26-NOV-1998) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Oct 21, 1998 this sequence version replaced gi:3723953. Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using PowerBlast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.





misc fe

/lab\_host="Top10F,"  
 /clone\_lib="SISNJ5s2"  
 /note="Organ: Stomach; Vector: pCNS; Site: 1: EcoRI;  
 Site 2: NotI; The po.y (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dT-selected mRNA by  
 priming with dT-tailed vector. The dT-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library. After analyzing and  
 sequencing about 2,000 ~ 3,000 colonies in original cDNA  
 library, the abundant cDNAs were selected and amplified by  
 PCR reaction using vector region primer including 7  
 promoter as 5' primer and N(dT)14 as 3' primer. The PCR  
 products were used as template for synthesis of  
 biotinylated single stranded RNA by in vitro transcription  
 reaction. The synthesized RNA probes were hybridized with  
 antisense single stranded cDNAs prepared from original  
 library and incubated with avidin-gel. After removing  
 DNA-RNA hybrids by centrifuge, the subtracted cDNA  
 libraries were constructed by transformation of the  
 remaining DNA into competent cells E. coli Top10F with  
 electroporation method."

## ORIGIN

Query Match 100.0%; Score 26; DB 12; Length 469;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGGTTGTACAGAC 26  
 |||  
 Db 192 TCAACTATGCGCGGTTGTACAGAC 217

## RESULT 19

BF819626/c  
 LOCUS BF819626 484 bp mRNA linear EST 13-JAN-2001  
 DEFINITION MR1-RT0028-101100-002-b04 RT0028 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF819626  
 VERSION BF819626.1 GI:12157598  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

AUTHORS  
 Dias Neto E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,  
 Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,  
 Goldman, G. F., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H.,  
 Brunstein, A., de Oliveira, P. S., Bucher, P., Jorgensen, C. V.,  
 O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and  
 Simpson, A. J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## MEDLINE

20202663

## PUBMED

10737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPBSP/LICR Human Cancer Genome  
 project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR1at2=MR1-RT0028-  
 101100-002-b04&f3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 484.

## FEATURES

Location/Qualifiers

source

1..484

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="RT0028"

/note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORSTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 26; DB 10; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTATGCGCGGTTGTACAGAC 26  
 |||  
 Db 447 TCAACTATGCGCGGTTGTACAGAC 422

## RESULT 20

BM694247

LOCUS BM694247 491 bp mRNA linear EST 28-FEB-2002

DEFINITION UI-E-CII-1-af0-1-16-0-UI-r1 UI-E-CII Homo sapiens cDNA clone

ACCESSION BM694247

VERSION BM694247.1 GI:19007505

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

Bonaldo, M. F., Lennon, G. and Soares, M. B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hagemar

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..491

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-CII-1-af0-1-16-0-UI"

/tissue\_type="RPE and Choroid"

/dev\_stage="adult"

/lab\_host="DHI08 (Life Technologies) (T1 phage resistant)"

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: [www-bio.lnl.gov/bbrp/image/image.html](http://www-bio.lnl.gov/bbrp/image/image.html)

Seq primer: -40UP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

## FEATURES

source

1..427  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2590179"  
/sex="female"  
/tissue\_type="papillary serous carcinoma"  
/lab\_host="DH105"  
/clone\_lib="NCI-CGAP\_Ov32"  
/note="Organ: ovary; Vector: pAMP1; mRNA made from ovarian carcinoma, cDNA made by oligo-dr priming.  
Non-directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Non-amplified library."

## ORIGIN

Query Match 100.0%; Score 26; DB 9; Length 427;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26

Db 324 TCAACTTATGCGCGTTGTACAGAC 299

## RESULT 17

BM767805

LOCUS

DEFINITION K-EST0050302 S1SNU5s2 Homo sapiens cDNA clone S1SNU5s2-12-E02 5', mRNA sequence.

ACCESSION BM767805

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 437)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.N., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 12 row: E column: 02

High quality sequence stop: 437.

Location/Qualifiers

1..437

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S1SNU5s2-12-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/lab\_host="Top10F"

/clone\_lib="S1SNU5s2"

/note="Organ: Stomach; Vector: pCMS; Site 1: EcoRI;

Site 2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including 3coor I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

## ORIGIN

Query Match 100.0%; Score 26; DB 12; Length 437;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAACTTATGCGCGTTGTACAGAC 26

Db 192 TCAACTTATGCGCGTTGTACAGAC 217

## RESULT 18

BM766631

LOCUS

DEFINITION K-EST0048574 S1SNU5s2 Homo sapiens cDNA clone S1SNU5s2-10-E02 5', mRNA sequence.

ACCESSION BM766631

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 469)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.N., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 10 row: E column: 02

High quality sequence stop: 469.

Location/Qualifiers

1..469

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S1SNU5s2-10-E02"

/sex="F"

/tissue\_type="Ascites"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-5"

```

ORIGIN
Query Match      100.0%; Score 26; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGGTTTGTACAGAC 26
    |||||
Db 101 TCAACTTATGCGCGGTTTGTACAGAC 126

RESULT 14
LOCUS C03540 422 bp mRNA linear EST 30-JUL-1996
DEFINITION C03540 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
3NHC1710, mRNA sequence.
ACCESSION C03540 GI:1466791
VERSION C03540.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
Nakamura,Y.
TITLE Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
JOURNAL Genomics 35 (1), 231-235 (1996)
MEDLINE 96299762
PUBMED 8661126
COMMENT Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
FEATURES
Source Location/Qualifiers
1..422
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="3NHC1710"
/dev_stage="adult"
/clone_lib="Human heart cDNA (Ynakamura)"
/notes="Organ: heart; normalized directionally cloned cDNA
from adult heart"

ORIGIN
Query Match      100.0%; Score 26; DB 13; Length 422;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGGTTTGTACAGAC 26
    |||||
Db 186 TCAACTTATGCGCGGTTTGTACAGAC 211

RESULT 15
LOCUS CD102141 423 bp mRNA linear EST 15-MAY-2003
DEFINITION AGENCOURT 14007825 NIH MGC 186 Homo sapiens cDNA clone
IMAGE:30370583 5', mRNA sequence.
ACCESSION CD102141
VERSION CD102141.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
cDNA Library Preparation: CLONTECH Laboratories, Inc.
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM159 row: m column: 24
High quality sequence stop: 417.
FEATURES
Source Location/Qualifiers
1..423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30370583"
/lab_host="DH10B (TL phage-resistant)"
/clone_lib="NIH_MGC_186"
/notes="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI
(ggcccattggcc); Site 2: SfiI (ggccgctggcc); Library is
oligo-dT primed and directionally cloned. cDNA was
prepared from a pooled samples of tissues from Skin,
meninges, duramater, pia matter and choroid plexus.
5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGCGCGGCGGCGGCGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library"
ORIGIN
Query Match      100.0%; Score 26; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAACTTATGCGCGGTTTGTACAGAC 26
    |||||
Db 202 TCAACTTATGCGCGGTTTGTACAGAC 227

RESULT 16
LOCUS AW085969/c 427 bp mRNA linear EST 14-OCT-1999
DEFINITION xc76d02.x1 NCI CGAP Ov32 Homo sapiens cDNA clone IMAGE:2590179 3',
similar to gb:R01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)
ALPHA CHAIN (HUMAN);, mRNA sequence.
ACCESSION AW085969
VERSION AW085969.1
KEYWORDS EST.
SOURCE GI:6041175
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Tissue Procurement: W. Marston Linahan, M.D., Rodrigo F. Chuaqui,
M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center

```

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 387)
AUTHORS    Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
TITLE      EST analysis of human adipose gene expression
JOURNAL    Unpublished (2002)
COMMENT     Contract: Gong Da-Wei
            Division of Endocrinology, Diabetes and Nutrition
            University of Maryland
            660 Redwood St., H4497, Baltimore, MD 21201, USA
            Tel: 410 706 1672
            Fax: 410 706 1622
            Email: dgong@medicine.umaryland.edu
PCR PRIMERS
FORWARD: CTCGGGAAGCGGCCATTGTGTGGT
BACKWARD: AATACGACTCACTATAGGCGGAATTGG
Seq primer: GTTGGTACCGGGAATTC.
FEATURES    Location/Qualifiers
            1..387
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /sex="Male and Female"
                /tissue_type="Adipose"
                /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
                /note="Vector: lambda triplex"
ORIGIN
Query Match      100.0%; Score 26; DB 14; Length 387;
Best Local Similarity 100.0%; Pred. NO. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 TCAACTTATGCGCGTTGTACAGAC 26
Db      142 TCAACTTATGCGCGTTGTACAGAC 167
RESULT 12
AA244273
LOCUS      AA244273
DEFINITION similar to gb:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)
ALPHA CHAIN (HUMAN); mRNA sequence.
ACCESSION AA244273
VERSION   AA244273.1 GI:1875023
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chusqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone Distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 673 Std Error: 0.00
            Seq primer: -28ml3 rev1 ET from Amersham
            High quality sequence stop: 371.
FEATURES    Location/Qualifiers
            1..405
                /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1007339"
/sex="Male"
/dev_stage="45 Years old"
/lab_host="DH10B"
/clone_lib="NCI-CGAP_Pri"
/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
ORIGIN
Query Match      100.0%; Score 26; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. NO. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 TCAACTTATGCGCGTTGTACAGAC 26
Db      145 TCAACTTATGCGCGTTGTACAGAC 170
RESULT 13
AA838010
LOCUS      AA838010
DEFINITION similar to gb:K01506 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DP(1)
ALPHA CHAIN (HUMAN); contains Alu repetitive element; mRNA
sequence.
ACCESSION AA838010
VERSION   AA838010.1 GI:2913667
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 435)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Stratagene, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone Distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 301.
FEATURES    Location/Qualifiers
            1..415
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:1418797"
                /sex="mixed"
                /tissue_type="colon tumor"
                /lab_host="SCLR (kanamycin resistant)"
                /clone_lib="NCI-CGAP Col2"
                /note="Organ: colon; Vector: Bluescript SK-; Site 1:
                EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
                Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'

```